(72) Inventors: FRANCIS-LANG, Helen; 1782 Pacific Avenue, Apt. 2, San Francisco, CA 94109 (US). FRIEDMAN, Lori; 113 Arundel Road, San Carlos, CA 94070 (US). BELVIN, Marcia; 921 Santa Fe Avenue, Albany, CA 94706 (US). PLOWMAN, Gregory, D.; 35 Winding Way, San Carlos, CA 94070 (US). LARSON, Jeffrey, S.; 1220 El Camino Real #305, Burlingame, CA 94010 (US). CHEN, Changyou; 6465 Crestwood Drive, Castro Valley, CA 94552 (US). ROBERSTON, Stephanie, A.; 255 Fowler Avenue, San Francisco, ca 94127 (US). LIOUBIN, Mario, N.; 3014 Los Prados, #A310, San Mateo, CA 94403 (US). SHI, Wen; 1324 8th Avenue, #4, San Francisco, CA 94112 (US). CHAN, Jocelyn; 1049 Cortland Avenue, San Francisco, CA 94110 (US). FUNKE, Roel, P.; 668, Sierra Point Road, Brisbane, CA 95005 (US). LI, Danxi; 90 Behr Avenue, #302, San Francisco, CA 94131 (US). KAUSELMANN, Gunther; Muenstereifeler Str. 39, 50937 Koeln (DE). TINTRUP, Hartmut; Waisenhausgasse 48, 50676 Cologne (DE). ZEVNIK, Branko; Gezelinallee 46, Leverkusen 51375 (DE). SCHOOR, Michael; Am Scherfenbrand 168, 51375 Leverkusen (DE). REARDON, Brian, John; 444 N. El Camino Real #101, San Mateo, CA 94401 (US).

(74) Agents: SHAYESTEH, Laleh et al.; Exelixis, Inc., P.O. Box 511, 170 Harbor Way, South San Francisco, CA 94083-0511 (US).

(54) Title: RRP SEQUENCES AND KNOCKOUT MICE AND USES THEREOF

(57) Abstract: RRP genes are identified as modulators of the p53 or p21 pathway, and thus are therapeutic targets for disorders associated with defective p53 or p21 function. Methods for identifying modulators of p53 or p21, comprising screening for agents that modulate the activity of RRP are provided. Modulating agents identified using the methods of the invention can be used to specifically inhibit growth of tumor cells that overexpress an RRP protein. mRRP1 knockout mice are also provided.

RRP SEQUENCES AND KNOCKOUT MICE AND USES THEREOF

REFERENCE TO RELATED APPLICATIONS

This application claims priority to US utility patent application 10/056,790 filed

5   1/23/2002, which is a Continuation In Part of US utility patent application 09/908,419

filed 7/18/2001, which claims priority to U.S. provisional patent applications 60/219,289,

filed 7/19/2000, 60/277,487, filed 3/21/2001, 60/277,471, filed 3/21/2001, 60/304,863,

filed 7/12/2001, 60/296,076 filed 6/5/2001, 60/305,017 filed 7/12/2001, 60/328,605 filed

10/10/2001, and 60/328,491 filed 10/10/2001. The contents of the prior applications are

10   hereby incorporated in their entirety.

BACKGROUND OF THE INVENTION

Signal transduction pathways are made up of growth factors, their receptors,

upstream regulators of the growth factors, and downstream intracellular kinase networks.

15   These pathways regulate many cellular processes, including proliferation, and appear to

play a key role in oncogenesis.

The epidermal growth factor receptor (EGFR) and its pathway members are among

the most widely explored signaling pathways. Signaling through this pathway elicits

diverse biological responses whose manifestations can include mitogenesis or apoptosis,

20   enhanced cell motility, protein secretion, and differentiation or dedifferentiation. Up-

regulated EGFR signaling has been implicated in organ morphogenesis, maintenance and

repair, and is correlated with invasion and metastasis of many types of tumors.

Upregulated EGF and EGFR in various tumor types leads to increased expression of

p21/Waf1/Cip1, leading to an arrest in the G1 phase of the cell cycle (Reddy KB et al.,

25   (1999) Int. J. Cancer 15:301-306).

Signal transduction pathways, such as the EGFR pathway, are evolutionarily

conserved among species as distant as the worm *Caenorhabditis elegans*, the fruit fly

Drosophila melanogaster, and vertebrates (Duffy JB, and Perrimon N, Curr. Opin. Cell

Biol. (1996) 8:231-238). In fact, ligands for the *Drosophila* EGFR (DER), known as Spitz

30   (Rutledge B, et al, Genes Dev. (1992) 6:1503-1517) and Gurken (Neuman-Silberberg FS,

and Schupbach T, Cell (1993) 75:165-174), are both similar to TGFα (transforming

growth factor alpha), the ligand for the vertebrate EGFR (Massague J, J Biol Chem.

(1990) 265:21393-21396). The rhomboid gene, which encodes a transmembrane protein,

is another upstream member of this pathway (Bier E., et al., Genes Dev. (1990) 4:190-

203). In *Drosophila*, rhomboid protein is a transmembrane serine protease, cleaves Spitz, and transforms Spitz from a membrane-bound to a secreted form, and thus triggers and upregulates the DER signaling pathway (Wasserman JD et al, Genes Devel (2000) 14:1651-1663; Lee JR, et al., (2001) Cell 107:161-171; Urban S, et al., (2001) Cell 107:

5 173-182). DNA sequences related to rhomboid have been identified in *C. elegans* (Wasserman JD, and Freeman M, Trends Cell Biol (1997) 7:431-436), and in mammals (Pascall JC, and Brown KD, FEBS letters (1998) 429:337-340; human: GI#3287191 and GI#7020534, among others; and rat: GI#3297936;), suggesting that rhomboid function may be evolutionarily conserved. Modulation of EGF-receptor related activity by

10 synthetic peptides or humanized monoclonal antibodies inhibit tumor growth (Baselga J, et al., (1998) Cancer Research 58:2825-2831; Park B-W, et al., (2000) Nature Biotechnology 18:194-197).

The p53 gene is mutated in over 50 different types of human cancers, including familial and spontaneous cancers, and is believed to be the most commonly mutated gene

15 in human cancer (Zambetti and Levine, FASEB (1993) 7:855-865; Hollstein, *et al.*, Nucleic Acids Res. (1994) 22:3551-3555). Greater than 90% of mutations in the p53 gene are missense mutations that alter a single amino acid that inactivates p53 function. Aberrant forms of human p53 are associated with poor prognosis, more aggressive tumors, metastasis, and short survival rates (Mitsudomi et al., Clin Cancer Res 2000 Oct;

20 6(10):4055-63; Koshland, Science (1993) 262:1953).

The human p53 protein normally functions as a central integrator of signals including DNA damage, hypoxia, nucleotide deprivation, and oncogene activation (Prives, Cell (1998) 95:5-8). In response to these signals, p53 protein levels are greatly increased with the result that the accumulated p53 activates cell cycle arrest (through activation of

25 p21/Waf1/Cip1) or apoptosis depending on the nature and strength of these signals. Indeed, multiple lines of experimental evidence have pointed to a key role for p53 as a tumor suppressor (Levine, Cell (1997) 88:323-331). For example, homozygous p53 "knockout" mice are developmentally normal but exhibit nearly 100% incidence of neoplasia in the first year of life (Donehower *et al.*, Nature (1992) 356:215-221).

30 The biochemical mechanisms and pathways through which p53 functions in normal and cancerous cells are not fully understood, but one clearly important aspect of p53 function is its activity as a gene-specific transcriptional activator. Among the genes with known p53-response elements are several with well-characterized roles in either

2

regulation of the cell cycle or apoptosis, including GADD45, p21/Waf1/Cip1, cyclin G, Bax, IGF-BP3, and MDM2 (Levine, Cell (1997) 88:323-331).

Modulating signal transduction pathway activity involved in tumor growth and development is essential in understanding the development of many cancers, and

5   eventually, for the treatment of cancer. The ability to manipulate the genomes of model organisms such as *Drosophila* provides a powerful means to analyze signal transduction pathwasy that, due to significant evolutionary conservation, have direct relevance to more complex vertebrate organisms. Due to a high level of gene and pathway conservation, the strong similarity of cellular processes, and the functional conservation of genes between

10   these model organisms and mammals, identification of the involvement of novel genes in particular pathways and their functions in such model organisms can directly contribute to the understanding of the correlative pathways and methods of modulating them in mammals (see, for example, Mechler BM et al., 1985 EMBO J 4:1551-1557; Gateff E. 1982 Adv. Cancer Res. 37: 33-74; Watson KL., et al., 1994 J Cell Sci. 18: 19-33; Miklos

15   GL, and Rubin GM. 1996 Cell 86:521-529; Wassarman DA, et al., 1995 Curr Opin Gen Dev 5: 44-50; and Booth DR. 1999 Cancer Metastasis Rev. 18: 261-284). For example, a genetic screen can be carried out in an invertebrate model organism having underexpression (e.g. knockout) or overexpression of a gene (referred to as a "genetic entry point") that yields a visible phenotype. Additional genes are mutated in a random or

20   targeted manner. When a gene mutation changes the original phenotype caused by the mutation in the genetic entry, the gene is identified as a "modifier" involved in the same or overlapping pathway as the genetic entry point. When the genetic entry point is an ortholog of a human gene implicated in a disease pathway, such as p53 or p21, modifier genes can be identified that may be attractive candidate targets for novel therapeutics.

25   All references cited herein, including sequence information in referenced Genbank identifier numbers and website references, are incorporated herein in their entireties.

SUMMARY OF THE INVENTION

We have discovered genes that modify the p53 or p21 pathway in *Drosophila*, and

30   identified their mammalian orthologs, hereinafter referred to as Rhomboid Related Proteins (RRP), and more specifically, RRP1 – RRP8, and mouse RRP1 (mRRP1). The invention provides isolated nucleic acid molecules that comprise nucleic acid sequences encoding RRP protein as well as fragments and derivatives thereof. Vectors and host cells comprising the RRP nucleic acid molecules are also described.

The invention provides methods for utilizing these p53 or p21 modifier genes and polypeptides to identify RRP modulating agents, which are candidate therapeutic agents that can be used in the treatment of disorders associated with defective p53 or p21 function.

In one embodiment, candidate p53 or p21 modulating agents are tested with an assay system comprising a RRP polypeptide or nucleic acid. Candidate agents that produce a change in the activity of the assay system relative to controls are identified as candidate p53 or p21 modulating agents. The assay system may be cell-based or cell-free. Candidate modulating agents include small molecule modulators, antibodies, and nucleic acid modulators. In one specific embodiment, a small molecule modulator is identified using a protease assay. In specific embodiments, the screening assay system is selected from an apoptosis assay, a cell proliferation assay, an angiogenesis assay, and a hypoxic induction assay.

In another embodiment, candidate p53 or p21 pathway modulating agents are further tested using a second assay system that detects changes in the p53 or p21 pathway, such as angiogenic, apoptotic, or cell proliferation changes produced by the originally identified candidate agent or an agent derived from the original agent. The second assay system may use cultured cells or non-human animals. In specific embodiments, the secondary assay system uses non-human animals, including animals predetermined to have a disease or disorder implicating the p53 or p21 pathway, such as an angiogenic, apoptotic, or cell proliferation disorder (e.g. cancer).

The invention further provides methods for modulating the p53 or p21 pathway in a mammalian cell by contacting the mammalian cell with an agent that specifically binds a RRP polypeptide or nucleic acid. The agent may be a small molecule modulator, a nucleic acid modulator, or an antibody and may be administered to a mammalian animal predetermined to have a pathology associated the p53 or p21 pathway.

Modulating agents identified using the methods of the invention can be used to specifically inhibit growth of tumor cells that overexpress an RRP protein.

The invention also provides transgenic knockout mice harboring disrupted RRP genes. The disruption may be heterozygous, leading to decreased expression of RRP, or homozygous, leading to lack of expression of the RRP gene. Cells from the mice as well as cells harboring disrupted RRP genes are also provided. Methods of producing antibody to RRP using the mice of the invention are also provided.

Targeting vectors to produce transgenic knockout mice are also provided.

Preferably, a targeting vector is provided that allows sequential deletion of vector

sequences from the same cell in the generation of the knockout mice.

5   **BRIEF DESCRIPTION OF THE DRAWINGS**

FIG.1 depicts the targeting vector, pEasyfloxFRTGK12 mRRP1.

FIG.2 shows the mRRP1 mouse targeting vector, the mRRP1 wild type (WT)

genomic region and the locations to which the vector is targeted, and the mRRP1 target

allele after homologous recombination (HR). Locations of the 5' external probe, the

10   internal probe, and the 3' external probe are underlined. The expected sizes of the

genomic fragments after digestion with the indicated restriction enzymes and Southern

hybridization are also indicated.

FIG.3 depicts the target allele after HR and after Cre induced deletion. The

expected sizes of the genomic fragments after digestion with restriction enzymes and

15   Southern hybridization are indicated for the target allele in each of the following states:

WT, HR, and Cre deleted.FIG.4 depicts the target allele after HR, after FLP induced

deletion, and after Cre induced deletion. The expected sizes of the genomic fragments

after digestion with the indicated restriction enzymes and Southern hybridization are

indicated for the target allele in each of the following states: WT, HR, FLP deleted, and

20   Cre deleted.

                          DETAILED DESCRIPTION OF THE INVENTION

Genetic screens were designed to identify modifiers of the p53 or p21 pathways in

Drosophila. Genetic modifier screens were carried out in which p53 (Ollmann M, et al.,

25   Cell 2000 101: 91-101) or p21 (Bourne HR, et al., Nature (1990) 348(6297):125-132;

Marshall CJ, Trends Genet (1991) 7(3):91-95) were overexpressed. *Drosophila* rhomboid

genes were identified as modifiers of the p53 or p21 pathways. Accordingly, vertebrate

orthologs of these modifiers, hereinafter referred to as RRP genes (i.e., nucleic acids and

polypeptides), are attractive drug targets for the treatment of pathologies associated with a

30   defective p53 or p21 signaling pathways, such as cancer. Further, gene targeting in mice

is an ideal method to investigate the function of a distinct protein in wild type and disease

states. In order to study the RRP1 function in mammals we generated the genomic

sequence of the RRP1 region in mice, deduced its cDNA and protein sequence, and then

produced targeted RRP1 knockout (KO) mice.

In vitro and in vivo methods of assessing RRP function as provided herein, and modulation of the RRP or their respective binding partners is useful for understanding the association of the p53 or p21 pathways and their members in normal and disease conditions and for developing diagnostics and therapeutic modalities for p53 or p21

5   related pathologies. RRP-modulating agents that act by inhibiting or enhancing RRP expression, directly or indirectly, for example, by affecting an RRP function such as enzymatic (e.g., catalytic) or binding activity, can be identified using methods provided herein. RRP-modulating agents include RRP related proteins (e.g. dominant negative mutants, and biotherapeutics); RRP-specific antibodies; RRP-specific antisense oligomers;

10   and chemical agents that specifically bind RRP or compete with RRP binding target. The invention provides methods of identifying and making RRP modulating agents, and their use in diagnosis, therapy and pharmaceutical development.

Preferred RRP-modulating agents specifically bind to RRP polypeptides and enhance or inhibit RRP function. Other preferred RRP-modulating agents are antisense

15   oligomers and RNAi that repress RRP gene expression or product activity by, for example, binding to and inhibiting the respective nucleic acid (i.e. DNA or mRNA). RRP-specific modulating agents may be evaluated by any convenient in vitro or in vivo assay for molecular interaction with an RRP polypeptide or nucleic acid.

The method of this invention is useful in the therapy of malignant or benign tumors

20   of mammals that overexpress RRP gene products.

Nucleic acids and polypeptides of the invention

Sequences related to RRP nucleic acids (RRP1: SEQ ID NO:1, RRP2: SEQ ID NO:3, RRP3: SEQID NO:5, RRP4: SEQ ID NO:7, RRP5: SEQ ID NO:9, RRP6: SEQ ID

25   NO:11, RRP7: SEQ ID NO:13, and RRP8: SEQ ID NO:15) and polypeptides (RRP1: SEQ ID NO:2, RRP2: SEQ ID NO: 4, RRP3: SEQID NO:6, RRP4: SEQ ID NO:8, RRP5: SEQ ID NO:10, RRP6: SEQ ID NO:12, RRP7: SEQ ID NO:14, and RRP8: SEQ ID NO:16) are available in the public databases (for RRP1: cDNA: GI#3287190, SEQ ID NO:18; proteins GI#3287191, SEQ ID NO:36; for RRP2: cDNAs: GI#s: 12762689 (SEQ ID

30   NO:19), 12096415 (SEQ ID NO:20), 6657080 (SEQ ID NO:21), 7947756 (SEQ ID NO:22), 3117010 (SEQ ID NO:23), 9339870 (SEQ ID NO:24), 7152496 (SEQ ID NO:25), 11317512 (SEQ ID NO:26), 11317513 (SEQ ID NO:27), and 11317511 (SEQ ID NO:28); proteins: GI#s:8923409 (SEQ ID NO:37), 12719522 (SEQ ID NO:38), and 11421817 (SEQ ID NO:39); for RRP3: cDNA: GI#10199673 (SEQ ID NO:29) and

GI#2003992 (SEQ ID NO:30); for RRP4: cDNA GI#11066249 (SEQ ID NO:31) and
protein GI#11066250 (SEQ ID NO:40); for RRP5: cDNA GI#11967982 (SEQ ID NO:32)
and protein GI#11967983 (SEQ ID NO:41); for RRP6: cDNA GI#10438685 (SEQ ID
NO:33) and protein GI#10438686 (SEQ ID NO:42); for RRP7: cDNA GI#10190733

5   . (SEQ ID NO:34) and protein GI#10190734 (SEQ ID NO:43); and for RRP8: cDNA
GI#11072100 (SEQ ID NO:35) and protein GI#11072101 (SEQ ID NO:44)). Sequences
of human and rat RRP1 were used to deduce the mouse RRP1 (mRRP1) cDNA (SEQ ID
NO:45), polypeptide (SEQ ID NO:46), and genomic (SEQ ID NO:47) sequences, as
described in Example VII. The mRRP1 cDNA sequence shares 69% identity with human

10 RRP1 and 88% identity with rat partial RRP1 for nucleotides 884-1340 of the mouse
RRP1. The mRRP1 protein shares 80% identity with human RRP1 and 88% identity with
rat partial RRP1 for amino acids 297-448 of mRRP1.

RRPs are a family of integral membrane proteins that contain five or more
transmembrane domains and three strongly conserved histidine residues in the putative

15 transmembrane regions. In a preferred embodiment, the invention provides RRP proteins
which comprise or consist of an amino acid sequence of SEQ ID NOs:4, 6, or 46, or
fragments or derivatives thereof.

The term "RRP polypeptide" refers to a full-length RRP protein or a functionally
active fragment or derivative thereof. A "functionally active" RRP fragment or derivative

20 exhibits one or more functional activities associated with a full-length, wild-type RRP
protein, such as antigenic or immunogenic activity, enzymatic activity, ability to bind
natural cellular substrates, etc. The functional activity of RRP proteins, derivatives and
fragments can be assayed by various methods known to one skilled in the art (Current
Protocols in Protein Science (1998) Coligan et al., eds., John Wiley & Sons, Inc.,

25 Somerset, New Jersey) and as further discussed below. For purposes herein, functionally
active fragments also include those fragments that comprise one or more structural
domains of an RRP, such as a protease or rhomboid domain or a binding domain.
Catalytic and other domains can be identified using the PFAM program (Bateman A., et
al., Nucleic Acids Res, 1999, 27:260-2; website at pfam.wustl.edu). Methods for

30 obtaining RRP polypeptides are also further described below. Preferred fragments are
functionally active, domain-containing fragments sharing at least 80% sequence identity or
similarity, preferably at least 85%, more preferably at least 90%, and most preferably at
least 95% sequence identity or similarity with a contiguous stretch of at least 25 amino
acids, preferably at least 50 amino acids, more preferably at least 100 amino acids, and in

some cases, the entire length of any one of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, and 46.
In further preferred embodiments, the fragment comprises the entire rhomboid domain
(PFAM 01694).

RRP protein derivatives typically share a certain degree of sequence identity or
5   sequence similarity with SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, or 46 or a fragment
thereof. RRP derivatives can be produced by various methods known in the art. The
manipulations which result in their production can occur at the gene or protein level. For
example, a cloned RRP gene sequence can be cleaved at appropriate sites with restriction
endonuclease(s) (Wells *et al.*, Philos. Trans. R. Soc. London SerA (1986) 317:415),
10   followed by further enzymatic modification if desired, isolated, and ligated *in vitro*, and
expressed to produce the desired derivative. Alternatively, an RRP gene can be mutated *in
vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination
sequences, or to create variations in coding regions and/or to form new restriction
endonuclease sites or destroy preexisting ones, to facilitate further *in vitro* modification.
15   A variety of mutagenesis techniques are known in the art such as chemical mutagenesis, *in
vitro* site-directed mutagenesis (Carter *et al.*, Nucl. Acids Res. (1986) 13:4331), use of
TAB® linkers (available from Pharmacia and Upjohn, Kalamazoo, MI), *etc.*

At the protein level, manipulations include post translational modification, *e.g.*
glycosylation, acetylation, phosphorylation, amidation, derivatization by known
20   protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other
cellular ligand, *etc.* Any of numerous chemical modifications may be carried out by
known technique (*e.g.* specific chemical cleavage by cyanogen bromide, trypsin,
chymotrypsin, papain, V8 protease, $NaBH_4$, acetylation, formylation, oxidation, reduction,
metabolic synthesis in the presence of tunicamycin, *etc.*). Derivative proteins can also be
25   chemically synthesized by use of a peptide synthesizer, for example to introduce
nonclassical amino acids or chemical amino acid analogs as substitutions or additions into
the RRP protein sequence.

Chimeric or fusion proteins can be made comprising an RRP protein or fragment
thereof (preferably comprising one or more structural or functional domains of the RRP
30   protein) joined at its amino- or carboxy-terminus via a peptide bond to an amino acid
sequence of a different protein. Chimeric proteins can be produced by any known method,
including: recombinant expression of a nucleic acid encoding the protein (comprising a
RRP-coding sequence joined in-frame to a coding sequence for a different protein);
ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences

to each other in the proper coding frame, and expressing the chimeric product; and protein synthetic techniques, *e.g.* by use of a peptide synthesizer.

The subject RRP polypeptides also encompass minor deletion mutants, including N-, and/or C-terminal truncations. Such deletion mutants are readily screened for RRP

5   competitive or dominant negative activity.

The term "RRP nucleic acid" refers to a DNA or RNA molecule that encodes a RRP polypeptide. In preferred embodiments, the nucleic acid encodes a polypeptide selected from the group consisting of SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, and 46. In some embodiments, the nucleic acid comprises a sequence selected from the group

10   consisting of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, and 45. In a specific embodiment, the invention provides an isolated nucleic acid which encodes a human RRP3 as shown in SEQ ID NO:5, and also an isolated nucleic acid that encodes a mouse RRP (mRRP1) as shown in SEQ ID NO:45.

The invention includes a fragment of a nucleic acid, such as a fragment that

15   encodes a binding domain of one of the full-length sequences of the invention. Fragments of an RRP nucleic acid sequence can be used for a variety of purposes. As an example, interfering RNA (RNAi) fragments, particularly double-stranded (ds) RNAi, can be used to generate loss-of-function phenotypes; which can, in turn, be used, among other uses, to determine gene function. Certain "antisense" fragments, i.e. that are reverse complements

20   of portions of the coding and/or untranslated regions (e.g. 5' UTR) have utility in inhibiting the function of RRP proteins. The fragments are of length sufficient to specifically hybridize with the corresponding RRP sequence. The fragments consist of or comprise at least 12, preferably at least 24, more preferably at least 36, and more preferably at least 96 contiguous nucleotides of RRP. When the fragments are flanked by

25   other nucleic acid sequences, the total length of the combined nucleic acid sequence is less than 15 kb, preferably less than 10 kb or less than 5kb, more preferably less than 2 kb, and in some cases, preferably less than 500 bases.

In other specific embodiments, preferred fragments of SEQ ID NO:5 encode extracellular or intracellular domains which are located at approximately nucleotides 248-

30   598, 665-796, 862-870, 934-943, 1006-1138, 1201-1225, and 1289-1336. Additional preferred fragments of SEQ ID NO:45 encode extracellular or intracellular domains which are located at approximately nucleotides 1-714, 774-912, 972-984, 1044-1089, 1149-1212, 1272-1305, 1365-1408. Preferred fragments may also include a binding domain or an RRP motif (e.g. PFAM 01694). These domains may be useful to locate the function

and/or binding partners of a protein. For example, a nucleic acid that encodes an extracellular or intracellular domain of a protein may be used to screen for binding partners related to the protein.

The subject nucleic acid sequences may consist solely of the RRP nucleic acid or fragments thereof. Alternatively, the subject nucleic acid sequences and fragments thereof may be joined to other components such as labels, peptides, agents that facilitate transport across cell membranes, hybridization-triggered cleavage agents or intercalating agents. The subject nucleic acid sequences and fragments thereof may also be joined to other nucleic acid sequences (i.e. they may comprise part of larger sequences) and are of synthetic/non-natural sequences and/or are isolated and/or are purified, i.e. unaccompanied by at least some of the material with which it is associated in its natural state. Preferably, the isolated nucleic acids constitute at least about 0.5%, and more preferably at least about 5% by weight of the total nucleic acid present in a given fraction, and are preferably recombinant, meaning that they comprise a non-natural sequence or a natural sequence joined to nucleotide(s) other than that which it is joined to on a natural chromosome.

The subject nucleic acids find a wide variety of applications including use as translatable transcripts, hybridization probes, PCR primers, diagnostic nucleic acids, etc.; use in detecting the presence of RRP genes and gene transcripts and in detecting or amplifying nucleic acids encoding additional RRP homologs and structural analogs. In diagnosis, RRP hybridization probes find use in identifying wild-type and mutant RRP alleles in clinical and laboratory samples. Mutant alleles are used to generate allele-specific oligonucleotide (ASO) probes for high-throughput clinical diagnoses. In therapy, therapeutic RRP nucleic acids are used to modulate cellular expression or intracellular concentration or availability of active RRP. In a preferred embodiment, the mouse RRP1 sequence is used to produce a targeting vector for production of mice that are deficient in RRP1 in a heterozygous or homozygous (i.e., knockout) manner.

In one preferred embodiment, the derivative nucleic acid encodes a polypeptide comprising a RRP3 amino acid sequence of SEQ ID NO:6, an mRRP1 amino acid sequence of SEQ ID NO:46, or a fragment or derivative thereof. A derivative RRP3 nucleic acid sequence, or fragment thereof, may comprise 100% sequence identity with SEQ ID NO:5 or 45, but be a derivative thereof in the sense that it has one or more modifications at the base or sugar moiety, or phosphate backbone. Examples of modifications are well known in the art (Bailey, Ullmann's Encyclopedia of Industrial

Chemistry (1998), 6th ed. Wiley and Sons). Such derivatives may be used to provide modified stability or any other desired property.

Preferably, the RRP polypeptide nucleic acid, fragment, ortholog, or derivative thereof has at least 70% sequence identity, preferably at least 80%, more preferably 85%,

5   still more preferably 90%, and most preferably at least 95% sequence identity with RRP. Normally, orthologs in different species retain the same function, due to presence of one or more protein motifs and/or 3-dimensional structures. As used herein, "percent (%) sequence identity" with respect to a subject sequence, or a specified portion of a subject sequence, is defined as the percentage of nucleotides or amino acids in the candidate

10   derivative sequence identical with the nucleotides or amino acids in the subject sequence (or specified portion thereof), after aligning the sequences and introducing gaps, if necessary to achieve the maximum percent sequence identity, as generated by the program WU-BLAST-2.0a19 (Altschul *et al.*, J. Mol. Biol. (1997) 215:403-410; website at blast.wustl.edu/blast/README.html) with all the search parameters set to default values.

15   The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched. A % identity value is determined by the number of matching identical nucleotides or amino acids divided by the sequence length for which the percent identity is being reported. "Percent

20   (%) amino acid sequence similarity" is determined by doing the same calculation as for determining % amino acid sequence identity, but including conservative amino acid substitutions in addition to identical amino acids in the computation.

A conservative amino acid substitution is one in which an amino acid is substituted for another amino acid having similar properties such that the folding or activity of the

25   protein is not significantly affected. Aromatic amino acids that can be substituted for each other are phenylalanine, tryptophan, and tyrosine; interchangeable hydrophobic amino acids are leucine, isoleucine, methionine, and valine; interchangeable polar amino acids are glutamine and asparagine; interchangeable basic amino acids are arginine, lysine and histidine; interchangeable acidic amino acids are aspartic acid and glutamic acid; and

30   interchangeable small amino acids are alanine, serine, threonine, cysteine and glycine.

Alternatively, an alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman (Smith and Waterman, 1981, Advances in Applied Mathematics 2:482-489; Smith and Waterman, 1981, J. of Molec.Biol., 147:195-197; Nicholas et al., 1998, "A Tutorial on Searching Sequence Databases and Sequence

Scoring Methods" (website at www.psc.edu) and references cited therein.; W.R. Pearson, 1991, Genomics 11:635-650). This algorithm can be applied to amino acid sequences by using the scoring matrix developed by Dayhoff (Dayhoff: Atlas of Protein Sequences and Structure, M. O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research

5   Foundation, Washington, D.C., USA), and normalized by Gribskov (Gribskov 1986 Nucl. Acids Res. 14(6):6745-6763). The Smith-Waterman algorithm is used to search databases for sequences similar to a query sequence. Smith-Waterman uses dynamic programming to determine how an optimal alignment between the query sequence and a database sequence can be produced. This alignment is obtained by determining what

10   transformations the query sequence would need to undergo to match the database sequence. Transformations include substituting one character for another and inserting or deleting a string of characters. A score is assigned for each character-to-character comparison--positive scores for exact matches and some substitutions, negative scores for other substitutions and insertions/deletions. The first character in an insertion or deletion

15   gap is scored with a gap open penalty and subsequent characters are scored with a gap extension penalty. Scores are obtained from statistically-derived scoring matrices. The combination of transformations that results in the highest score is used to generate an alignment between the query sequence and database sequence. Smith-Waterman algorithm may be employed where default parameters are used for scoring (for example,

20   gap open penalty of 12, gap extension penalty of two). From the data generated the "Match" value reflects "sequence identity."

Derivative nucleic acid molecules of the subject nucleic acid molecules include sequences that hybridize to the nucleic acid sequence of SEQ ID NOs:1, 3, 5, 7, 9, 1, 13, 15, or 45. The stringency of hybridization can be controlled by temperature, ionic

25   strength, pH, and the presence of denaturing agents such as formamide during hybridization and washing. Conditions routinely used are set out in readily available procedure texts (*e.g.*, Current Protocol in Molecular Biology, Vol. 1, Chap. 2.10, John Wiley & Sons, Publishers (1994); Sambrook *et al.*, Molecular Cloning, Cold Spring Harbor (1989)). In some embodiments, a nucleic acid molecule of the invention is capable

30   of hybridizing to a nucleic acid molecule containing the nucleotide sequence of any one of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, or 45 under stringent hybridization conditions that comprise: prehybridization of filters containing nucleic acid for 8 hours to overnight at 65° C in a solution comprising 6X single strength citrate (SSC) (1X SSC is 0.15 M NaCl, 0.015 M Na citrate; pH 7.0), 5X Denhardt's solution, 0.05% sodium pyrophosphate and

100 μg/ml herring sperm DNA; hybridization for 18-20 hours at 65° C in a solution containing 6X SSC, 1X Denhardt's solution, 100 μg/ml yeast tRNA and 0.05% sodium pyrophosphate; and washing of filters at 65° C for 1h in a solution containing 0.2X SSC and 0.1% SDS (sodium dodecyl sulfate).

5   In other embodiments, moderately stringent hybridization conditions are used that comprise: pretreatment of filters containing nucleic acid for 6 h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH7.5), 5mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 μg/ml denatured salmon sperm DNA; hybridization for 18-20h at 40° C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl

10   (pH7.5), 5mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 μg/ml salmon sperm DNA, and 10% (wt/vol) dextran sulfate; followed by washing twice for 1 hour at 55° C in a solution containing 2X SSC and 0.1% SDS.

          Alternatively, low stringency conditions can be used that comprise: incubation for 8 hours to overnight at 37° C in a solution comprising 20% formamide, 5 x SSC, 50 mM

15   sodium phosphate (pH 7.6), 5X Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured sheared salmon sperm DNA; hybridization in the same buffer for 18 to 20 hours; and washing of filters in 1 x SSC at about 37° C for 1 hour.

Isolation, Production, Expression, and Mis-expression of RRP Nucleic Acids and

20   ## Polypeptides

          RRP nucleic acids and polypeptides, useful for identifying and testing agents that modulate RRP function and for other applications related to the involvement of RRP in the p53 or p21 pathways. RRP nucleic acids and derivatives and orthologs thereof may be obtained using any available method. For instance, techniques for isolating cDNA or

25   genomic DNA sequences of interest by screening DNA libraries or by using polymerase chain reaction (PCR) are well known in the art. In general, the particular use for the protein will dictate the particulars of expression, production, and purification methods. For instance, production of proteins for use in screening for modulating agents may require methods that preserve specific biological activities of these proteins, whereas

30   production of proteins for antibody generation may require structural integrity of particular epitopes. Expression of proteins to be purified for screening or antibody production may require the addition of specific tags (e.g., generation of fusion proteins). Overexpression of an RRP protein for assays used to assess RRP function, such as involvement in cell cycle regulation or hypoxic response, may require expression in eukaryotic cell lines

capable of these cellular activities. Techniques for the expression, production, and purification of proteins are well known in the art; any suitable means therefore may be used (e.g., Higgins SJ and Hames BD (eds.) Protein Expression: A Practical Approach, Oxford University Press Inc., New York 1999; Stanbury PF et al., Principles of

5   Fermentation Technology, 2nd edition, Elsevier Science, New York, 1995; Doonan S (ed.) Protein Purification Protocols, Humana Press, New Jersey, 1996; Coligan JE et al, Current Protocols in Protein Science (eds.), 1999, John Wiley & Sons, New York; U.S. Pat. No. 6,165,992). In particular embodiments, recombinant RRP is expressed in a cell line known to have defective p53 or p21 function (e.g. for p53: SAOS-2 osteoblasts, H1299

10   lung cancer cells, C33A and HT3 cervical cancer cells, HT-29 and DLD-1 colon cancer cells, among others, and for p21: HCT116 colon cancer cells, among others, available from American Type Culture Collection (ATCC), Manassas, VA). The recombinant cells are used in cell-based screening assay systems of the invention, as described further below.

15   The nucleotide sequence encoding an RRP polypeptide can be inserted into any appropriate expression vector. The necessary transcriptional and translational signals, including promoter/enhancer element, can derive from the native RRP gene and/or its flanking regions or can be heterologous. A variety of host-vector expression systems may be utilized, such as mammalian cell systems infected with virus (e.g. vaccinia virus,

20   adenovirus, etc.); insect cell systems infected with virus (e.g. baculovirus); microorganisms such as yeast containing yeast vectors, or bacteria transformed with bacteriophage, plasmid, or cosmid DNA. A host cell strain that modulates the expression of, modifies, and/or specifically processes the gene product may be used.

           To detect expression of the RRP gene product, the expression vector can comprise

25   a promoter operably linked to an RRP gene nucleic acid, one or more origins of replication, and, one or more selectable markers (e.g. thymidine kinase activity, resistance to antibiotics, etc.). Alternatively, recombinant expression vectors can be identified by assaying for the expression of the RRP gene product based on the physical or functional properties of the RRP protein in in vitro assay systems (e.g. immunoassays).

30   The RRP protein, fragment, or derivative may be optionally expressed as a fusion, or chimeric protein product (i.e. it is joined via a peptide bond to a heterologous protein sequence of a different protein), for example to facilitate purification or detection. A chimeric product can be made by ligating the appropriate nucleic acid sequences encoding the desired amino acid sequences to each other using standard methods and expressing the

14

chimeric product. A chimeric product may also be made by protein synthetic techniques, *e.g.* by use of a peptide synthesizer (Hunkapiller *et al.*, Nature (1984) 310:105-111).

Once a recombinant cell that expresses the RRP gene sequence is identified, the gene product can be isolated and purified using standard methods (*e.g.* ion exchange,

5   affinity, and gel exclusion chromatography; centrifugation; differential solubility; electrophoresis, cite purification reference). Alternatively, native RRP proteins can be purified from natural sources, by standard methods (*e.g.* immunoaffinity purification). Once a protein is obtained, it may be quantified and its activity measured by appropriate methods, such as immunoassay, bioassay, or other measurements of physical properties,

10   such as crystallography.

The methods of this invention may also use cells that have been engineered for altered expression (mis-expression) of RRP or other genes associated with the p53 or p21 pathway. As used herein, mis-expression encompasses ectopic expression, over-expression, under-expression, and non-expression (*e.g.* by gene knock-out or blocking

15   expression that would otherwise normally occur).

Genetically modified animals

Animal models that have been genetically modified to alter RRP expression may be used in *in vivo* assays to test for activity of a candidate p53 or p21 modulating agent, or

20   to further assess the role of RRP in a p53 or p21 pathway process such as apoptosis or cell proliferation. Preferably, the altered RRP expression results in a detectable phenotype, such as decreased or increased levels of cell proliferation, angiogenesis, or apoptosis compared to control animals having normal RRP expression. The genetically modified animal may additionally have altered p53 or p21 expression (e.g. p53 or p21 knockout).

25   Preferred genetically modified animals are mammals such as primates, rodents (preferably mice), cows, horses, goats, sheep, pigs, dogs and cats. Preferred non-mammalian species include zebrafish, *C. elegans*, and *Drosophila*. Preferred genetically modified animals are transgenic animals having a heterologous nucleic acid sequence present as an extrachromosomal element in a portion of its cells, i.e. mosaic animals (see, for example,

30   techniques described by Jakobovits, 1994, Curr. Biol. 4:761-763.) or stably integrated into its germ line DNA (i.e., in the genomic sequence of most or all of its cells). Heterologous nucleic acid is introduced into the germ line of such transgenic animals by genetic manipulation of, for example, embryos or embryonic stem cells of the host animal.

Methods of making transgenic animals are well-known in the art (for transgenic mice see Brinster et al., Proc. Nat. Acad. Sci. USA 82: 4438-4442 (1985), U.S. Pat. Nos. 4,736,866 and 4,870,009, both by Leder et al., U.S. Pat. No. 4,873,191 by Wagner et al., U.S. Pat. No. 6,127,598, by German et al., and Hogan, B., Manipulating the Mouse

5   Embryo, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., (1986); for particle bombardment see U.S. Pat. No., 4,945,050, by Sandford et al.; for transgenic Drosophila see Rubin and Spradling, Science (1982) 218:348-53 and U.S. Pat. No. 4,670,388; for transgenic insects see Berghammer A.J. et al., A Universal Marker for Transgenic Insects (1999) Nature 402:370-371; for transgenic Zebrafish see Lin S.,

10   Transgenic Zebrafish, Methods Mol Biol. (2000);136:375-3830); for microinjection procedures for fish, amphibian eggs and birds see Houdebine and Chourrout, Experientia (1991) 47:897-905; for transgenic rats see Hammer et al., Cell (1990) 63:1099-1112; and for culturing of embryonic stem (ES) cells and the subsequent production of transgenic animals by the introduction of DNA into ES cells using methods such as electroporation,

15   calcium phosphate/DNA precipitation and direct injection see, e.g., Teratocarcinomas and Embryonic Stem Cells, A Practical Approach, E. J. Robertson, ed., IRL Press (1987)). Clones of the nonhuman transgenic animals can be produced according to available methods (see Wilmut, I. et al. (1997) Nature 385:810-813; and PCT International Publication Nos. WO 97/07668 and WO 97/07669).

20   In one embodiment, the transgenic animal is a "knock-out" animal having a heterozygous or homozygous alteration in the sequence of an endogenous RRP gene that results in a decrease of RRP function, preferably such that RRP expression is undetectable or insignificant. Knock-out animals are typically generated by homologous recombination with a vector comprising a transgene having at least a portion of the gene to be knocked

25   out. Typically a deletion, addition or substitution has been introduced into the transgene to functionally disrupt it. The transgene can be a human gene (e.g., from a human genomic clone) but more preferably is an ortholog of the human gene derived from the transgenic host species. For example, a mouse RRP gene is used to construct a homologous recombination vector suitable for altering an endogenous RRP gene in the

30   mouse genome as shown in Example VIII. Detailed methodologies for homologous recombination in mice are available (see Capecchi, Science (1989) 244:1288-1292; Joyner et al., Nature (1989) 338:153-156). Procedures for the production of non-rodent transgenic mammals and other animals are also available (Houdebine and Chourrout, supra; Pursel et al., Science (1989) 244:1281-1288; Simms et al., Bio/Technology (1988)

16

6:179-183). In a preferred embodiment, knock-out animals, such as mice harboring a knockout of a specific gene, may be used to produce antibodies against the human counterpart of the gene that has been knocked out (Claesson MH et al. (1994) Scand J Immunol 40:257-264; Declerck PJ, et al., (1995) J Biol Chem 270:8397-8400).

5   In another embodiment, the transgenic animal is a "knock-in" animal having an alteration in its genome that results in altered expression (e.g., increased (including ectopic) or decreased expression) of the RRP gene, e.g., by introduction of additional copies of RRP, or by operatively inserting a regulatory sequence that provides for altered expression of an endogenous copy of the RRP gene. Such regulatory sequences include

10   inducible, tissue-specific, and constitutive promoters and enhancer elements. The knock-in can be homozygous or heterozygous.

Transgenic nonhuman animals can also be produced that contain selected systems allowing for regulated expression of the transgene. One example of such a system that may be produced is the cre/loxP recombinase system of bacteriophage P1 (Lakso *et al.*,

15   PNAS (1992) 89:6232-6236; U.S. Pat. No. 4,959,317). If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other

20   containing a transgene encoding a recombinase. Another example of a recombinase system is the FLP recombinase system of Saccharomyces cerevisiae (O'Gorman et al. (1991) Science 251:1351-1355; U.S. Pat. No. 5,654,182). In a preferred embodiment, both Cre-LoxP and Flp-Frt are used in the same system to regulate expression of the transgene, and for sequential deletion of vector sequences in the same cell, such as shown

25   in Example VIII.

The genetically modified animals can be used in genetic studies to further elucidate the p53 or p21 pathway, as animal models of disease and disorders implicating defective p53 or p21 function, and for *in vivo* testing of candidate therapeutic agents, such as those identified in the screens described below. Gene targeting in mice is an ideal method to

30   investigate the function of a distinct protein in wild type and disease states. Further, animal models deficient in rhomboid RRP sequence and function are desirable tools for modulating the EGFR signaling pathway, for testing the effect of candidate compounds against RRP, and for production of antibodies against human RRP, among others. The candidate therapeutic agents are administered to a genetically modified animal having

17

altered RRP function and phenotypic changes are compared with appropriate control animals such as genetically modified animals that receive placebo treatment, and/or animals with unaltered RRP expression that receive candidate therapeutic agent.

In additional to the above-described genetically modified animals having altered
5   RRP function, animal models having defective p53 or p21 function (and otherwise normal RRP function), can be used in the methods of the present invention. For example, a p53 or p21 knockout mouse can be used to assess, *in vivo*, the activity of a candidate p53 or p21 modulatory agent identified in one of the *in vitro* assays described below. p53 or p21 knockout mice are described in the literature (p53: Jacks et al., Nature 2001;410:1111-
10 1116, 1043-1044; Donehower *et al.*, supra; p21:Umanoff H, et al., Proc Natl Acad Sci U S A 1995 Feb 28;92(5):1709-13).

Modulating Agents

The invention provides methods to identify agents that interact with and/or
15 modulate the function of RRP and/or the p53 or p21 pathway. Such agents are useful in a variety of diagnostic and therapeutic applications associated with the p53 or p21 pathways, as well as in further analysis of the RRP protein and its contribution to the p53 or p21 pathways. Accordingly, the invention also provides methods for modulating the p53 or p21 pathway comprising the step of specifically modulating RRP activity by administering
20 an RRP-interacting or -modulating agent.

In a preferred embodiment, RRP-modulating agents inhibit or enhance RRP activity or otherwise affect normal RRP function, including transcription, protein expression, protein localization, and cellular or extra-cellular activity. In a further preferred embodiment, the candidate p53 or p21 pathway- modulating agent specifically
25 modulates the function of the RRP. The phrases "specific modulating agent", "specifically modulates", etc., are used herein to refer to modulating agents that directly bind to the RRP polypeptide or nucleic acid, and preferably inhibit, enhance, or otherwise alter, the function of the RRP. The term also encompasses modulating agents that alter the interaction of the RRP with a binding partner or substrate (e.g. by binding to a binding
30 partner of an RRP, or to a protein/binding partner complex, and inhibiting function).

Preferred RRP-modulating agents include small molecule compounds; RRP-interacting proteins, including antibodies and other biotherapeutics; antisense and RNA inhibitors. The modulating agents may be formulated in pharmaceutical compositions, for example, as compositions that may comprise other active ingredients, as in combination

therapy, and/or suitable carriers or excipients. Techniques for formulation and administration of the compounds may be found in "Remington's Pharmaceutical Sciences" Mack Publishing Co., Easton, PA, 19th edition.

5   **Small molecule modulators**

Small molecules, are often preferred to modulate function of proteins with enzymatic function, and/or containing protein interaction domains. Chemical agents, referred to in the art as "small molecule" compounds are typically organic, non-peptide molecules, having a molecular weight less than 10,000, preferably less than 5,000, more

10   preferably less than 1,000, and most preferably less than 500. This class of modulators includes chemically synthesized molecules, for instance, compounds from combinatorial chemical libraries. Synthetic compounds may be rationally designed or identified based on known or inferred properties of the RRP protein or may be identified by screening compound libraries. Alternative appropriate modulators of this class are natural products,

15   particularly secondary metabolites from organisms such as plants or fungi, which can also be identified by screening compound libraries for RRP–modulating activity. Methods for generating and obtaining compounds are well known in the art (Schreiber SL, Science (2000) 151: 1964-1969; Radmann J and Gunther J, Science (2000) 151:1947-1948).

Small molecule modulators identified from screening assays, as described below,

20   can be used as lead compounds from which candidate clinical compounds may be designed, optimized, and synthesized. Such clinical compounds may have utility in treating pathologies associated with the p53 or p21 pathway. The activity of candidate small molecule modulating agents may be improved several-fold through iterative secondary functional validation, as further described below, structure determination, and

25   candidate modulator modification and testing. Additionally, candidate clinical compounds are generated with specific regard to clinical and pharmacological properties. For example, the reagents may be derivatized and re-screened using *in vitro* and *in vivo* assays to optimize activity and minimize toxicity for pharmaceutical development.

30   **Protein Modulators**

Specific RRP-interacting proteins are useful in a variety of diagnostic and therapeutic applications related to the p53 or p21 pathway and related disorders, as well as in validation assays for other RRP-modulating agents. In a preferred embodiment, RRP-interacting proteins affect normal RRP function, including transcription, protein

expression, protein localization, and cellular or extra-cellular activity. In another embodiment, RRP-interacting proteins are useful in detecting and providing information about the function of RRP proteins, as is relevant to p53 or p21 related disorders, such as cancer (e.g., for diagnostic means).

5   An RRP-interacting protein may be endogenous, i.e. one that naturally interacts genetically or biochemically with an RRP, such as TGFα, EGF, amphiregulin, heregulin, a member of the RRP pathway that modulates RRP expression, localization, and/or activity. RRP-modulators include dominant negative forms of RRP-interacting proteins and of RRP proteins themselves. Yeast two-hybrid and variant screens offer preferred methods for

10   identifying endogenous RRP-interacting proteins (Finley, R. L. et al. (1996) in DNA Cloning-Expression Systems: A Practical Approach, eds. Glover D. & Hames B. D (Oxford University Press, Oxford, England), pp. 169-203; Fashema SF et al., Gene (2000) 250:1-14; Drees BL Curr Opin Chem Biol (1999) 3:64-70; Vidal M and Legrain P Nucleic Acids Res (1999) 27:919-29; and U.S. Pat. No. 5,928,868). Mass spectrometry is an

15   alternative preferred method for the elucidation of protein complexes (reviewed in, e.g., Pandley A and Mann M, Nature (2000) 405:837-846; Yates JR 3rd, Trends Genet (2000) 16:5-8).

        An RRP-interacting protein may be an exogenous protein, such as an RRP-specific antibody or a T-cell antigen receptor (see, e.g., Harlow and Lane (1988) Antibodies, A

20   Laboratory Manual, Cold Spring Harbor Laboratory). RRP antibodies are further discussed below.

        In preferred embodiments, an RRP-interacting protein specifically binds an RRP protein. In alternative preferred embodiments an RRP-modulating agent binds an RRP substrate, binding partner, or cofactor.

25

        Antibodies

        In another embodiment, the protein modulator is an RRP specific antibody agonist or antagonist. The antibodies have therapeutic and diagnostic utilities, and can be used in screening assays to identify RRP modulators. For example, uses for antibodies include the

30   detection of an RRP protein in a biological sample and the inhibition of RRP activity, for instance, to block the development of an oncogenic disorder. The antibodies can also be used in dissecting the portions of the RRP pathway responsible for various cellular responses and in the general processing and maturation of the RRP.

Antibodies that specifically bind RRP polypeptides can be generated using known methods. Preferably the antibody is specific to a mammalian ortholog of RRP polypeptide, and more preferably, to human RRP. Antibodies may be polyclonal, monoclonal (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab

5   fragments, F(ab').sub.2 fragments, fragments produced by a FAb expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Monoclonal antibodies with affinities of 10^8 M^{-1} preferably 10^9 M^{-1} to 10^{10} M^{-1}, or stronger can be made by standard procedures as described (Harlow and Lane, Antibodies: A Laboratory Manual, CSH Laboratory (1988); Goding (1986) Monoclonal Antibodies:

10   Principles and Practice (2d ed) Academic Press, New York; and U.S. Pat. Nos. 4,381,292; 4,451,570; and 4,618,577). Antibodies may be generated against crude cell extracts of RRP or substantially purified fragments thereof. If RRP fragments are used, they preferably comprise at least 10, and more preferably, at least 20 contiguous amino acids of an RRP protein. In a particular embodiment, RRP-specific antigens and/or immunogens

15   are coupled to carrier proteins that stimulate the immune response. For example, the subject polypeptides are covalently coupled to the keyhole limpet hemocyanin (KLH) carrier, and the conjugate is emulsified in Freund's complete adjuvant, which enhances the immune response. An appropriate immune system such as a laboratory rabbit or mouse is immunized according to conventional protocols. In a preferred embodiment, due to close

20   similarity of RRP sequences from mice and humans, transgenic mice that are RRP deficient or RRP knockout, such as those generated in the present invention (Example VIII), are used to produce antibodies against human RRP.

The presence of RRP-specific antibodies is assayed by an appropriate assay such as a solid phase enzyme-linked immunosorbant assay (ELISA) using immobilized

25   corresponding RRP polypeptides. Other assays, such as radioimmunoassays or fluorescent assays might also be used.

Chimeric antibodies specific to RRP polypeptides can be made that contain different portions from different animal species. For instance, a human immunoglobulin constant region may be linked to a variable region of a murine mAb, such that the

30   antibody derives its biological activity from the human antibody, and its binding specificity from the murine fragment. Chimeric antibodies are produced by splicing together genes that encode the appropriate regions from each species (Morrison et al., Proc. Natl. Acad. Sci. (1984) 81:6851-6855; Neuberger et al., Nature (1984) 312:604-608; Takeda et al., Nature (1985) 31:452-454). Humanized antibodies, which are a form of

chimeric antibodies, can be generated by grafting complementary-determining regions (CDRs) (Carlos, T. M., J. M. Harlan. 1994. Blood 84:2068-2101) of mouse antibodies into a background of human framework regions and constant regions by recombinant DNA technology (Riechmann LM, et al., 1988 Nature 323: 323-327). Humanized

5   antibodies contain ~10% murine sequences and ~90% human sequences, and thus further reduce or eliminate immunogenicity, while retaining the antibody specificities (Co MS, and Queen C. 1991 Nature 351: 501-501; Morrison SL. 1992 Ann. Rev. Immun. 10:239-265). Humanized antibodies and methods of their production are well-known in the art (U.S. Pat. No. 5,530,101; U.S. Pat. No. 5,585,089; U.S. Pat. No. 5,693,762, and

10   U.S. Pat. No. 6,180,370).

RRP-specific single chain antibodies which are recombinant, single chain polypeptides formed by linking the heavy and light chain fragments of the Fv regions via an amino acid bridge, can be produced by methods known in the art (U.S. Pat. No. 4,946,778; Bird, Science (1988) 242:423-426; Huston et al., Proc. Natl. Acad. Sci. USA

15   (1988) 85:5879-5883; and Ward et al., Nature (1989) 334:544-546).

Other suitable techniques for antibody production involve in vitro exposure of lymphocytes to the antigenic polypeptides or alternatively to selection of libraries of antibodies in phage or similar vectors (Huse et al., Science (1989) 246:1275-1281). As used herein, T-cell antigen receptors are included within the scope of antibody modulators

20   (Harlow and Lane, 1988, *supra*).

The polypeptides and antibodies of the present invention may be used with or without modification. Frequently, antibodies will be labeled by joining, either covalently or non-covalently, a substance that provides for a detectable signal, or that is toxic to cells that express the targeted protein (Menard S, et al., Int J. Biol Markers (1989) 4:131-134).

25   A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, fluorescent emitting lanthanide metals, chemiluminescent moieties, bioluminescent moieties, magnetic particles, and the like (U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350;

30   3,996,345; 4,277,437; 4,275,149; and 4,366,241). Also, recombinant immunoglobulins may be produced (U.S. Pat. No. 4,816,567). Antibodies to cytoplasmic polypeptides may be delivered and reach their targets by conjugation with membrane-penetrating toxin proteins (U.S. Pat. No. 6,086,900).

When used therapeutically in a patient, the antibodies of the subject invention are typically administered parenterally, when possible at the target site, or intravenously. The therapeutically effective dose and dosage regimen is determined by clinical studies. Typically, the amount of antibody administered is in the range of about 0.1 mg/kg –to
5   about 10 mg/kg of patient weight. For parenteral administration, the antibodies are formulated in a unit dosage injectable form (e.g., solution, suspension, emulsion) in association with a pharmaceutically acceptable vehicle. Such vehicles are inherently nontoxic and non-therapeutic. Examples are water, saline, Ringer's solution, dextrose solution, and 5% human serum albumin. Nonaqueous vehicles such as fixed oils, ethyl
10   oleate, or liposome carriers may also be used. The vehicle may contain minor amounts of additives, such as buffers and preservatives, which enhance isotonicity and chemical stability or otherwise enhance therapeutic potential. The antibodies' concentrations in such vehicles are typically in the range of about 1 mg/ml–to about10 mg/ml. Immunotherapeutic methods are further described in the literature (US Pat. No. 5,859,206;
15   WO0073469).

The selection of an appropriate antibody subclass for therapy will depend upon the nature of the tumor antigen. For example, an IgM may be preferred when the antigen is highly specific for the tumor target and rarely occurs on normal cells. However, the IgG subclass may be preferred when the tumor-associated antigen is also expressed in normal
20   tissues, even at much lower levels. The binding of at least two IgG molecules in close proximity is required to activate complement, a serum protein that combines with antibodies to form a defense against cellular antigens. The normal tissues that express smaller amounts of the antigen and bind fewer IgG molecules may thus incur less complement-mediated damage. Furthermore, since IgGs are smaller than IgMs, they may
25   more readily localize to tumor tissue.

Immune responses may assist in the delivery or efficacy of an anti-tumor treatment. There is evidence that complement activation leads to an inflammatory response and macrophage activation (Uananue and Benecerraf, Textbook of Immunology, 2nd Edition, Williams & Wilkins, p. 218 (1984)). Activated macrophages more
30   preferentially destroy tumor cells than normal cells (Fidler and Poste, Springer Semin. Immunopathol. 5, 161 (1982)). Also, the increased vasodilation accompanying inflammation may increase the ability of anti-cancer agents, such as chemotherapeutic drugs or radiolabeled antibodies to localize in tumors. While a significant detriment of standard chemotherapy or radiation treatment is damage to healthy cells, the antigen-

antibody combinations specified by this invention may circumvent many of the problems normally caused by the heterogeneity of tumor cell populations. Additionally, purified antigens (Hakomori, Ann. Rev. Immunol. (1984) 2:103) or the related anti-idiotypic antibodies (Nepom et al., Proc. Natl. Acad. Sci, (1985) 81:2864; Koprowski et al., Proc.

5   Natl. Acad. Sci. (1984) 81:216) which recognize the hypervariance among the same epitopes in different individuals could be used to induce an active immune response in human cancer patients. Such a response includes the formation of antibodies capable of activating human complement and mediating antibody-dependent cell-mediated cytotoxicity and by such mechanisms cause tumor destruction.

10

Specific biotherapeutics

In a preferred embodiment, an RRP-interacting protein may have biotherapeutic applications. Biotherapeutic agents formulated in pharmaceutically acceptable carriers and dosages may be used to activate or inhibit signal transduction pathways. This

15   modulation may be accomplished by binding a ligand, thus inhibiting the activity of the pathway; or by binding a receptor, either to inhibit activation of, or to activate, the receptor. Alternatively, the biotherapeutic may itself be a ligand capable of activating or inhibiting a receptor. Biotherapeutic agents and methods of producing them are described in detail in U.S. Pat. No. 6,146,628.

20   Since RRP is a receptor, its ligand(s), antibodies to the ligand(s) or the RRP itself may be used as biotherapeutics to modulate the activity of RRP in the p53 or p21 pathway.

Nucleic Acid Modulators

Other preferred RRP-modulating agents comprise nucleic acid molecules, such as

25   antisense oligomers or double stranded RNA (dsRNA), which generally inhibit RRP activity.

Preferred antisense oligomers that interfere with the function of the RRP nucleic acid such as DNA replication, transcription, translocation of the RRP RNA to the site of protein translation, translation of protein from the RRP RNA, splicing of the RRP RNA to

30   yield one or more mRNA species, and catalytic activity which may be engaged in or facilitated by the RRP RNA. Double-stranded RNA inhibition (dsRNAi) is another preferred RRP-modulating agent. For convenience, the term "antisense modulator", as used herein, includes antisense oligomers and dsRNAi.

In one embodiment, the antisense oligomer is an oligonucleotide that is sufficiently complementary to an RRP mRNA to bind to and prevent translation, preferably by binding to the 5' untranslated region. RRP-specific antisense oligonucleotides, preferably range from at least 6 to about 200 nucleotides. In some embodiments the oligonucleotide is

5   preferably at least 10, 15, or 20 nucleotides in length. In other embodiments, the oligonucleotide is preferably less than 50, 40, or 30 nucleotides in length. The oligonucleotide can be DNA or RNA or a chimeric mixture or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone. The oligonucleotide may

10   include other appending groups such as peptides, agents that facilitate transport across the cell membrane, hybridization-triggered cleavage agents, and intercalating agents.

In another embodiment, the antisense oligomer is a phosphothioate morpholino oligomer (PMO). PMOs are assembled from four different morpholino subunits, each of which contain one of four genetic bases (A, C, G, or T) linked to a six-membered

15   morpholine ring. Polymers of these subunits are joined by non-ionic phosphodiamidate intersubunit linkages. Details of how to make and use PMOs and other antisense oligomers are well known in the art (e.g. see WO99/18193; Probst JC, Antisense Oligodeoxynucleotide and Ribozyme Design, Methods. (2000) 22(3):271-281; Summerton J, and Weller D. 1997 Antisense Nucleic Acid Drug Dev. ;7:187-95, US Pat. No.

20   5,235,033; and US Pat No. 5,378,841).

Antisense oligomers are commonly used as research reagents, diagnostics, and therapeutics. For example, antisense oligonucleotides, which are able to inhibit gene expression with exquisite specificity, are often used to elucidate the function of particular genes (see, for example, U.S. Pat. No. 6,165,790). Antisense oligomers are also used, for

25   example, to distinguish between functions of various members of a biological pathway. Antisense oligomers have been employed as therapeutic moieties in the treatment of disease states in animals and man and have been demonstrated in numerous clinical trials to be safe and effective (Milligan JF, et al, Current Concepts in Antisense Drug Design, J Med Chem. (1993) 36:1923-1937; Tonkinson JL et al., Antisense Oligodeoxynucleotides

30   as Clinical Therapeutic Agents, Cancer Invest. (1996) 14:54-65). Accordingly, in one aspect of the invention, an RRP-specific antisense oligomer is used in an assay to further elucidate the role of the RRP in the p53 or p21 pathway, and/or its relationship to other members of the pathway. In another aspect of the invention, an RRP-specific antisense oligomer is used as a therapeutic agent for treatment of p53 or p21-related disease states.

Alternative preferred RRP-modulating agents are double-stranded RNA species mediating RNA interference (RNAi). RNAi is the process of sequence-specific, post-transcriptional gene silencing in animals and plants, initiated by double-stranded RNA (dsRNA) that is homologous in sequence to the silenced gene. Methods relating to the use

5   of RNAi to silence genes in *C. elegans*, *Drosophila*, plants, and humans are known in the art (Fire A, et al., 1998   Nature 391:806-811; Fire, A. Trends Genet. 15, 358-363 (1999); Sharp, P. A. RNA interference 2001. Genes Dev. 15, 485-490 (2001); Hammond, S. M., et al., Nature Rev. Genet. 2, 110-1119 (2001); Tuschl, T. Chem. Biochem. 2, 239-245 (2001); Hamilton, A. et al., Science 286, 950-952 (1999); Hammond, S. M., et al.,

10   Nature 404, 293-296 (2000); Zamore, P. D., et al., Cell 101, 25-33 (2000); Bernstein, E., et al., Nature 409, 363-366 (2001); Elbashir, S. M., et al., Genes Dev. 15, 188-200 (2001); WO0129058; WO9932619; Elbashir SM, et al., 2001 Nature 411:494-498).

Assay Systems

15   The invention provides assay systems for identifying specific modulators of RRP activity. As used herein, an "assay system" encompasses all the components required for performing and analyzing results of an assay that detects and/or measures a particular event. In general, primary assays are used to identify or confirm a modulator's specific biochemical or molecular effect with respect to the RRP nucleic acid or protein. In

20   general, secondary assays further assess the activity of a RRP modulating agent identified by a primary assay and may confirm that the modulating agent affects RRP in a manner relevant to the p53 or p21 pathway. In some cases, RRP modulators will be directly tested in a secondary assay.

        In a preferred embodiment, the screening method comprises contacting a suitable

25   assay system comprising an RRP polypeptide with a candidate agent under conditions whereby, but for the presence of the agent, the system provides a reference activity (e.g. protease activity), which is based on the particular molecular event the screening method detects. A statistically significant difference between the agent-biased activity and the reference activity indicates that the candidate agent modulates RRP activity, and hence the

30   p53 or p21 pathway.

Primary Assays

The type of modulator tested generally determines the type of primary assay.

Primary assays for small molecule modulators

For small molecule modulators, screening assays are used to identify candidate modulators. Screening assays may be cell-based or may use a cell-free system that recreates or retains the relevant biochemical reaction of the target protein (reviewed in

5   Sittampalam GS *et al.*, Curr Opin Chem Biol (1997) 1:384-91 and accompanying references). As used herein the term "cell-based" refers to assays using live cells, dead cells, or a particular cellular fraction, such as a membrane, endoplasmic reticulum, or mitochondrial fraction. The term "cell free" encompasses assays using substantially purified protein (either endogenous or recombinantly produced), partially purified cellular

10 extracts, or crude cellular extracts. Screening assays may detect a variety of molecular events, including protein-DNA interactions, protein-protein interactions (*e.g.*, receptor-ligand binding), transcriptional activity (*e.g.*, using a reporter gene), enzymatic activity (*e.g.*, via a property of the substrate), activity of second messengers, immunogenicty and changes in cellular morphology or other cellular characteristics. Appropriate screening

15 assays may use a wide range of detection methods including fluorescent, radioactive, colorimetric, spectrophotometric, and amperometric methods, to provide a read-out for the particular molecular event detected.

In a preferred embodiment, screening assays uses fluorescence technologies, including fluorescence polarization, time-resolved fluorescence, and fluorescence

20 resonance energy transfer. These systems offer means to monitor protein-protein or DNA-protein interactions in which the intensity of the signal emitted from dye-labeled molecules depends upon their interactions with partner molecules (*e.g.*, Selvin PR, Nat Struct Biol (2000) 7:730-4; Fernandes PB, Curr Opin Chem Biol (1998) 2:597-603; Hertzberg RP and Pope AJ, Curr Opin Chem Biol (2000) 4:445-451).

25   Cell-based screening assays usually require systems for recombinant expression of RRP and any auxiliary proteins demanded by the particular assay. Cell-free assays often use recombinantly produced purified or substantially purified proteins. Appropriate methods for generating recombinant proteins produce sufficient quantities of proteins that retain their relevant biological activities and are of sufficient purity to optimize activity

30 and assure assay reproducibility. Yeast two-hybrid and variant screens, and mass spectrometry provide preferred methods for determining protein-protein interactions and elucidation of protein complexes. In certain applications, when RRP-interacting proteins are used in screens to identify small molecule modulators, the binding specificity of the interacting protein to the RRP protein may be assayed by various known methods such as

27

substrate processing (e.g. ability of the candidate RRP-specific binding agents to function as negative effectors in RRP-expressing cells), binding equilibrium constants (usually at least about $10^7 M^{-1}$, preferably at least about $10^8 M^{-1}$, more preferably at least about $10^9 M^{-1}$), and immunogenicity (e.g. ability to elicit RRP specific antibody in a heterologous host

5   such as a mouse, rat, goat or rabbit). For enzymes and receptors, binding may be assayed by, respectively, substrate and ligand processing.

The screening assay may measure a candidate agent's ability to specifically bind to or modulate activity of a RRP polypeptide, a fusion protein thereof, or to cells or membranes bearing the polypeptide or fusion protein. The RRP polypeptide can be full

10   length or a fragment thereof that retains functional RRP activity. The RRP polypeptide may be fused to another polypeptide, such as a peptide tag for detection or anchoring, or to another tag. The RRP polypeptide is preferably human or mouse RRP, or is an ortholog or derivative thereof as described above. In a preferred embodiment, the screening assay detects candidate agent-based modulation of RRP interaction with a binding target, such as

15   an endogenous or exogenous protein or other substrate that has RRP –specific binding activity, and can be used to assess normal RRP gene function.

Suitable assay formats that may be adapted to screen for RRP modulators are known in the art. Preferred screening assays are high throughput or ultra high throughput and thus provide automated, cost-effective means of screening compound libraries for lead

20   compounds (Fernandes PB, 1998, *supra*; Sundberg SA, Curr Opin Biotechnol 2000, 11:47-53).

A variety of suitable assay systems may be used to identify candidate RRP and p53 or p21 pathway modulators (e.g. U.S. Pat. No. 6,020,135 (p53 modulation), U.S. Pat. No. 6,114,132 (phosphatase and protease assays)). Specific preferred assays are described in

25   more detail below.

Protease Assays. Proteases are enzymes that cleave protein substrates at specific sites. Exemplary assays detect the alterations in the spectral properties of an artificial substrate that occur upon protease-mediated cleavage. In one example, synthetic caspase

30   substrates containing four amino acid proteolysis recognition sequences, separating two different fluorescent tags are employed; fluorescence resonance energy transfer detects the proximity of these fluorophores, which indicates whether the substrate is cleaved (Mahajan NP *et al.*, Chem Biol (1999) 6:401-409).

Endogenous protease inhibitors may inhibit protease activity. In an example of an assay developed for either proteases or protease inhibitors, a biotinylated substrate is coated on a titer plate and hydrolyzed with the protease; the unhydrolyzed substrate is quantified by reaction with alkaline phosphatase-streptavidin complex and detection of the

5   reaction product. The activity of protease inhibitors correlates with the activity of the alkaline phosophatase indicator enzyme (Gan Z *et al.*, Anal Biochem 1999) 268:151-156).

Apoptosis assays. Assays for apoptosis may be performed by terminal deoxynucleotidyl transferase-mediated digoxigenin-11-dUTP nick end labeling (TUNEL)

10   assay. The TUNEL assay is used to measure nuclear DNA fragmentation characteristic of apoptosis (Lazebnik *et al.*, 1994, Nature 371, 346), by following the incorporation of fluorescein-dUTP (Yonehara *et al.*, 1989, J. Exp. Med. 169, 1747). Apoptosis may further be assayed by acridine orange staining of tissue culture cells (Lucas, R., et al., 1998, Blood 15:4730-41). An apoptosis assay system may comprise a cell that expresses an RRP, and

15   that optionally has defective p53 or p21 function (e.g. p53 or p21 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the apoptosis assay system and changes in induction of apoptosis relative to controls where no test agent is added, identify candidate p53 or p21 modulating agents. In some embodiments of the invention, an apoptosis assay may be used as a secondary assay to test a candidate p53 or

20   p21 modulating agents that is initially identified using a cell-free assay system. An apoptosis assay may also be used to test whether RRP function plays a direct role in apoptosis. For example, an apoptosis assay may be performed on cells that over- or under-express RRP relative to wild type cells. Differences in apoptotic response compared to wild type cells suggests that the RRP plays a direct role in the apoptotic response.

25   Apoptosis assays are described further in US Pat. No. 6,133,437.

Cell proliferation and cell cycle assays. Cell proliferation may be assayed via bromodeoxyuridine (BRDU) incorporation. This assay identifies a cell population undergoing DNA synthesis by incorporation of BRDU into newly-synthesized DNA.

30   Newly-synthesized DNA may then be detected using an anti-BRDU antibody (Hoshino *et al.*, 1986, Int. J. Cancer 38, 369; Campana *et al.*, 1988, J. Immunol. Meth. 107, 79), or by other means.

Cell Proliferation may also be examined using $[^3H]$-thymidine incorporation (Chen, J., 1996, Oncogene 13:1395-403; Jeoung, J., 1995, J. Biol. Chem. 270:18367-73).

This assay allows for quantitative characterization of S-phase DNA syntheses. In this assay, cells synthesizing DNA will incorporate [^{3}H]-thymidine into newly synthesized DNA. Incorporation can then be measured by standard techniques such as by counting of radioisotope in a scintillation counter (e.g., Beckman LS 3800 Liquid Scintillation

5   Counter).

Cell proliferation may also be assayed by colony formation in soft agar (Sambrook et al., Molecular Cloning, Cold Spring Harbor (1989)). For example, cells transformed with RRP are seeded in soft agar plates, and colonies are measured and counted after two weeks incubation.

10   Involvement of a gene in the cell cycle may be assayed by flow cytometry. Cells transfected with an RRP may be stained with propidium iodide and evaluated in a flow cytometer (available from Becton Dickinson).

Accordingly, a cell proliferation or cell cycle assay system may comprise a cell that expresses an RRP, and that optionally has defective p53 or p21 function (e.g. p53 or

15   p21 is over-expressed or under-expressed relative to wild-type cells). A test agent can be added to the assay system and changes in cell proliferation or cell cycle relative to controls where no test agent is added, identify candidate p53 or p21 modulating agents. In some embodiments of the invention, the cell proliferation or cell cycle assay may be used as a secondary assay to test a candidate p53 or p21 modulating agents that is initially identified

20   using another assay system such as a cell-free kinase assay system. A cell proliferation assay may also be used to test whether RRP function plays a direct role in cell proliferation or cell cycle. For example, a cell proliferation or cell cycle assay may be performed on cells that over- or under-express RRP relative to wild type cells. Differences in proliferation or cell cycle compared to wild type cells suggests that the RRP

25   plays a direct role in cell proliferation or cell cycle. A cell proliferation assay may also be used to identify candidate agents that modulate cell proliferation. For example, cells that have decreased expression of RRP or that are RRP knockouts, such as mouse cells generated in the present invention (Example VIII) are treated with candidate agents. Changes in cell proliferation relative to control cells where no agent is added indicate that

30   the candidate agent modulates cell proliferation.

Angiogenesis. Angiogenesis may be assayed using various human endothelial cell systems, such as umbilical vein, coronary artery, or dermal cells. Suitable assays include Alamar Blue based assays (available from Biosource International) to measure

proliferation; migration assays using fluorescent molecules, such as the use of Becton
Dickinson Falcon HTS FluoroBlock cell culture inserts to measure migration of cells
through membranes in presence or absence of angiogenesis enhancer or suppressors; and
tubule formation assays based on the formation of tubular structures by endothelial cells

5   on Matrigel® (Becton Dickinson). Accordingly, an angiogenesis assay system may
comprise a cell that expresses an RRP, and that optionally has defective p53 or p21
function (e.g. p53 or p21 is over-expressed or under-expressed relative to wild-type cells).
A test agent can be added to the angiogenesis assay system and changes in angiogenesis
relative to controls where no test agent is added, identify candidate p53 or p21 modulating

10   agents. In some embodiments of the invention, the angiogenesis assay may be used as a
secondary assay to test a candidate p53 or p21 modulating agents that is initially identified
using another assay system. An angiogenesis assay may also be used to test whether RRP
function plays a direct role in cell proliferation. For example, an angiogenesis assay may
be performed on cells that over- or under-express RRP relative to wild type cells.

15   Differences in angiogenesis compared to wild type cells suggests that the RRP plays a
direct role in angiogenesis.

         Hypoxic induction. The alpha subunit of the transcription factor, hypoxia
inducible factor-1 (HIF-1), is upregulated in tumor cells following exposure to hypoxia in

20   vitro. Under hypoxic conditions, HIF-1 stimulates the expression of genes known to be
important in tumor cell survival, such as those encoding glyolytic enzymes and VEGF.
Induction of such genes by hypoxic conditions may be assayed by growing cells
transfected with RRP in hypoxic conditions (such as with 0.1% O_2, 5% CO_2, and balance
N_2, generated in a Napco 7001 incubator (Precision Scientific)) and normoxic conditions,

25   followed by assessment of gene activity or expression by Taqman®. For example, a
hypoxic induction assay system may comprise a cell that expresses an RRP, and that
optionally has a mutated p53 or p21 (e.g. p53 or p21 is over-expressed or under-expressed
relative to wild-type cells). A test agent can be added to the hypoxic induction assay
system and changes in hypoxic response relative to controls where no test agent is added,

30   identify candidate p53 or p21 modulating agents. In some embodiments of the invention,
the hypoxic induction assay may be used as a secondary assay to test a candidate p53 or
p21 modulating agents that is initially identified using another assay system. A hypoxic
induction assay may also be used to test whether RRP function plays a direct role in the
hypoxic response. For example, a hypoxic induction assay may be performed on cells that

over- or under-express RRP relative to wild type cells. Differences in hypoxic response compared to wild type cells suggest that the RRP plays a direct role in hypoxic induction.

Cell adhesion. Cell adhesion assays measure adhesion of cells to purified
5   adhesion proteins, or adhesion of cells to each other, in presence or absence of candidate modulating agents.

Cell-protein adhesion assays measure the ability of agents to modulate the adhesion of cells to purified proteins. For example, recombinant proteins are produced, diluted to 2.5g/mL in PBS, and used to coat the wells of a microtiter plate. The wells used
10   for negative control are not coated. Coated wells are then washed, blocked with 1% BSA, and washed again. Compounds are diluted to 2x final test concentration and added to the blocked, coated wells. Cells are then added to the wells, and the unbound cells are washed off. Retained cells are labeled directly on the plate by adding a membrane-permeable fluorescent dye, such as calcein-AM, and the signal is quantified in a fluorescent
15   microplate reader.

Cell-cell adhesion assays measure the ability of agents to modulate binding of cell adhesion proteins with their native ligands. These assays use cells that naturally or recombinantly express the adhesion protein of choice. In an exemplary assay, cells expressing the cell adhesion protein are plated in wells of a multiwell plate. Cells
20   expressing the ligand are labeled with a membrane-permeable fluorescent dye, such as BCECF , and allowed to adhere to the monolayers in the presence of candidate agents. Unbound cells are washed off, and bound cells are detected using a fluorescence plate reader.

High-throughput cell adhesion assays have also been described. In one such assay,
25   small molecule ligands and peptides are bound to the surface of microscope slides using a microarray spotter, intact cells are then contacted with the slides, and unbound cells are washed off. In this assay, not only the binding specificity of the peptides and modulators against cell lines are determined, but also the functional cell signaling of attached cells using immunofluorescence techniques in situ on the microchip is measured (Falsey JR et
30   al., Bioconjug Chem. 2001 May-Jun;12(3):346-53).

Certain screening assays may also be used to test antibody and nucleic acid modulators; for nucleic acid modulators, appropriate assay systems involve RRP mRNA expression.

Primary assays for antibody modulators

For antibody modulators, appropriate primary assays test is a binding assay that tests the antibody's affinity to and specificity for the RRP protein. Methods for testing antibody affinity and specificity are well known in the art (Harlow and Lane, 1988, 1999,

5   *supra*). The enzyme-linked immunosorbant assay (ELISA) is a preferred method for detecting RRP-specific antibodies; others include FACS assays, radioimmunoassays, and fluorescent assays.

Primary assays for nucleic acid modulators

10   For nucleic acid modulators, primary assays may test the ability of the nucleic acid modulator to inhibit or enhance RRP gene expression, preferably mRNA expression. In general, expression analysis comprises comparing RRP expression in like populations of cells (*e.g.*, two pools of cells that endogenously or recombinantly express RRP) in the presence and absence of the nucleic acid modulator. Methods for analyzing mRNA and

15   protein expression are well known in the art. For instance, Northern blotting, slot blotting, ribonuclease protection, quantitative RT-PCR (*e.g.*, using the TaqMan®, PE Applied Biosystems), or microarray analysis may be used to confirm that RRP mRNA expression is reduced in cells treated with the nucleic acid modulator (*e.g.*, Current Protocols in Molecular Biology (1994) Ausubel FM *et al.*, *eds.*, John Wiley & Sons, Inc., chapter 4;

20   Freeman WM *et al.*, Biotechniques (1999) 26:112-125; Kallioniemi OP, Ann Med 2001, 33:142-147; Blohm DH and Guiseppi-Elie, A Curr Opin Biotechnol 2001, 12:41-47). Protein expression may also be monitored. Proteins are most commonly detected with specific antibodies or antisera directed against either the RRP protein or specific peptides. A variety of means including Western blotting, ELISA, or in situ detection, are available

25   (Harlow E and Lane D, 1988 and 1999, *supra*).

Secondary Assays

Secondary assays may be used to further assess the activity of RRP-modulating agent identified by any of the above methods to confirm that the modulating agent affects

30   RRP in a manner relevant to the p53 or p21 pathway. As used herein, RRP-modulating agents encompass candidate clinical compounds or other agents derived from previously identified modulating agent. Secondary assays can also be used to test the activity of a modulator on a particular genetic or biochemical pathway or to test the specificity of the modulator's interaction with RRP.

33

Secondary assays generally compare like populations of cells or animals (*e.g.*, two pools of cells or animals that endogenously or recombinantly express RRP) in the presence and absence of the candidate modulator. In general, such assays test whether treatment of cells or animals with a candidate RRP–modulating agent results in changes in the p53 or

5   p21 pathway in comparison to untreated (or mock- or placebo-treated) cells or animals. Certain assays use "sensitized genetic backgrounds", which, as used herein, describe cells or animals engineered for altered expression of genes in the p53 or p21 or interacting pathways.

10   *Cell-based assays*

Cell based assays may use a variety of mammalian cell lines known to have defective p53 or p21 function (e.g. for p53: SAOS-2 osteoblasts, H1299 lung cancer cells, C33A and HT3 cervical cancer cells, HT-29 and DLD-1 colon cancer cells, among others, and for p21: HCT116 colon cancer cells, among others, available from American Type

15   Culture Collection (ATCC), Manassas, VA). Cell based assays may detect endogenous p53 or p21 pathway activity or may rely on recombinant expression of p53 or p21 pathway components. Any of the aforementioned assays may be used in this cell-based format. Candidate modulators are typically added to the cell media but may also be injected into cells or delivered by any other efficacious means.

20

Animal Assays

A variety of non-human animal models of normal or defective p53 or p21 pathway may be used to test candidate RRP modulators. Models for defective p53 or p21 pathway typically use genetically modified animals that have been engineered to mis-express (*e.g.*,

25   over-express or lack expression in) genes involved in the p53 or p21 pathway. Assays generally require systemic delivery of the candidate modulators, such as by oral administration, injection, etc.

In a preferred embodiment, p53 or p21 pathway activity is assessed by monitoring neovascularization and angiogenesis. Animal models with defective and normal p53 or

30   p21 are used to test the candidate modulator's affect on RRP in Matrigel® assays. Matrigel® is an extract of basement membrane proteins, and is composed primarily of laminin, collagen IV, and heparin sulfate proteoglycan. It is provided as a sterile liquid at 4° C, but rapidly forms a solid gel at 37° C. Liquid Matrigel® is mixed with various angiogenic agents, such as bFGF and VEGF, or with human tumor cells which over-

express the RRP. The mixture is then injected subcutaneously into the female athymic
nude mice (Taconic, Germantown, NY) to support an intense vascular response. Mice
with Matrigel® pellets may be dosed via oral (PO), intraperitoneal (IP), or intravenous
(IV) routes with the candidate modulator. Mice are euthanized 5 - 12 days post-injection,
5   and the Matrigel® pellet is harvested for hemoglobin analysis (Sigma plasma hemoglobin
kit). Hemoglobin content of the gel is found to correlate the degree of neovascularization
in the gel.

In another preferred embodiment, the effect of the candidate modulator on RRP is
assessed via tumorigenicity assays. In one example, xenograft human tumors are
10   implanted subcutaneously (SC) into female athumic nude mice, 6-7 week old, as single
cell suspensions either from a pre-existing tumor or from *in vitro* culture. The tumors
which express the RRP endogenously are injected in the flank, 1×10^5 to 1×10^7 cells per
mouse in a volume of 100 μL using a 27gauge needle. Mice are then ear tagged and
tumors are measured twice weekly. Candidate modulator treatment is initiated on the day
15   the mean tumor weight reaches 100 mg. Candidate modulator is delivered intravenously
(IV), subcutaneously (SC), intraperitoneously (IP), or orally (PO) by bolus administration.
Depending upon the pharmacokinetics (PK) of each unique candidate modulator, dosing
can be performed multiple times per day. The tumor weight is assessed by measuring
perpendicular diameters with a caliper and calculated by multiplying the measurements of
20   diameters in two dimensions. At the end of the experiment, the excised tumors maybe
utilized for biomarker identification or further analyses. For immunohistochemistry
staining, xenograft tumors are fixed in 4% paraformaldehyde, 0.1M phosphate, PH 7.2, for
6 hours at 4°C, immersed in 30% sucrose in PBS, and rapidly frozen in isopetane cooled
with liquid nitrogen.

25

Diagnostic and therapeutic uses

Specific RRP-modulating agents are useful in a variety of diagnostic and
therapeutic applications where disease or disease prognosis is related to defects in the p53
or p21 pathway, such as angiogenic, apoptotic, or cell proliferation disorders.
30   Accordingly, the invention also provides methods for modulating the p53 or p21 pathway
in a cell, preferably a cell pre-determined to have defective p53 or p21 function,
comprising the step of administering an agent to the cell that specifically modulates RRP
activity.

The discovery that RRP is implicated in p53 or p21 pathway provides for a variety of methods that can be employed for the diagnostic and prognostic evaluation of diseases and disorders involving defects in the p53 or p21 pathway and for the identification of subjects having a predisposition to such diseases and disorders.

5   Various expression analysis methods can be used to diagnose whether RRP expression occurs in a particular sample, including Northern blotting, slot blotting, ribonuclease protection, quantitative RT-PCR, and microarray analysis. (e.g., Current Protocols in Molecular Biology (1994) Ausubel FM et al., eds., John Wiley & Sons, Inc., chapter 4; Freeman WM et al., Biotechniques (1999) 26:112-125; Kallioniemi OP, Ann

10   Med 2001, 33:142-147; Blohm and Guiseppi-Elie, Curr Opin Biotechnol 2001, 12:41-47). Tissues having a disease or disorder implicating defective p53 or p21 signaling that express an RRP, are identified as amenable to treatment with an RRP modulating agent. In a preferred application, the p53 or p21 defective tissue overexpresses an RRP relative to normal tissue. For example, a Northern blot analysis of mRNA from tumor and normal

15   cell lines, or from tumor and matching normal tissue samples from the same patient, using full or partial RRP cDNA sequences as probes, can determine whether particular tumors express or overexpress RRP. Alternatively, the TaqMan® is used for quantitative RT-PCR analysis of RRP expression in cell lines, normal tissues and tumor samples (PE Applied Biosystems).

20   Various other diagnostic methods may be performed, for example, utilizing reagents such as the RRP oligonucleotides, and antibodies directed against an RRP, as described above for: (1) the detection of the presence of RRP gene mutations, or the detection of either over- or under-expression of RRP mRNA relative to the non-disorder state; (2) the detection of either an over- or an under-abundance of RRP gene product

25   relative to the non-disorder state; and (3) the detection of perturbations or abnormalities in the signal transduction pathway mediated by RRP.

Thus, in a specific embodiment, the invention is drawn to a method for diagnosing a disease in a patient, the method comprising: a) obtaining a biological sample from the patient; b) contacting the sample with a probe for RRP expression; c) comparing results

30   from step (b) with a control; and d) determining whether step (c) indicates a likelihood of disease. Preferably, the disease is cancer, most preferably a cancer as shown in TABLE 1. The probe may be either DNA or protein, including an antibody.

EXAMPLES

The following experimental section and examples are offered by way of illustration
and not by way of limitation.

5   I. Drosophila p53 and p21 screens

The Drosophila p53 gene was overexpressed specifically in the wing using the
vestigial margin quadrant enhancer. Increasing quantities of Drosophila p53 (titrated
using different strength transgenic inserts in 1 or 2 copies) caused deterioration of normal
wing morphology from mild to strong, with phenotypes including disruption of pattern and
10   polarity of wing hairs, shortening and thickening of wing veins, progressive crumpling of
the wing and appearance of dark "death" inclusions in wing blade. In a screen designed to
identify enhancers and suppressors of Drosophila p53, homozygous females carrying two
copies of p53 were crossed to 5663 males carrying random insertions of a piggyBac
transposon (Fraser M et al., Virology (1985) 145:356-361). Progeny containing insertions
15   were compared to non-insertion-bearing sibling progeny for enhancement or suppression
of the p53 phenotypes. Sequence information surrounding the piggyBac insertion site was
used to identify the modifier genes. Modifiers of the wing phenotype were identified as
members of the p53 pathway.

An overexpression screen was carried out in *Drosophila* to identify genes that
20   interact with the cyclin dependent kinase inhibitor, p21 (Bourne HR, et al., Nature (1990)
348(6297):125-132; Marshall CJ, Trends Genet (1991) 7(3):91-95). Expression of the p21
gene in the eye causes deterioration of normal eye morphology. Modifiers of the eye
phenotype were identified as members of the p21 pathway.

25   II. High-Throughput In Vitro Fluorescence Polarization Assay

Fluorescently-labeled RRP peptide/substrate are added to each well of a 96-well
microtiter plate, along with a test agent in a test buffer (10 mM HEPES, 10 mM NaCl, 6
mM magnesium chloride, pH 7.6). Changes in fluorescence polarization, determined by
using a Fluorolite FPM-2 Fluorescence Polarization Microtiter System (Dynatech
30   Laboratories, Inc), relative to control values indicates the test compound is a candidate
modifier of RRP activity.

III. High-Throughput In Vitro Binding Assay.

^{33}P-labeled RRP peptide is added in an assay buffer (100 mM KCl, 20 mM HEPES pH 7.6, 1 mM MgCl$_2$, 1% glycerol, 0.5% NP-40, 50 mM beta-mercaptoethanol, 1 mg/ml BSA, cocktail of protease inhibitors) along with a test agent to the wells of a Neutralite-

5   avidin coated assay plate and incubated at 25°C for 1 hour. Biotinylated substrate is then added to each well and incubated for 1 hour. Reactions are stopped by washing with PBS, and counted in a scintillation counter. Test agents that cause a difference in activity relative to control without test agent are identified as candidate p53 or p21 modulating agents.

10

IV. Immunoprecipitations and Immunoblotting

For coprecipitation of transfected proteins, 3×10^6 appropriate recombinant cells containing the RRP proteins are plated on 10-cm dishes and transfected on the following day with expression constructs. The total amount of DNA is kept constant in each

15   transfection by adding empty vector. After 24 h, cells are collected, washed once with phosphate-buffered saline and lysed for 20 min on ice in 1 ml of lysis buffer containing 50 mM Hepes, pH 7.9, 250 mM NaCl, 20 mM -glycerophosphate, 1 mM sodium orthovanadate, 5 mM p-nitrophenyl phosphate, 2 mM dithiothreitol, protease inhibitors (complete, Roche Molecular Biochemicals), and 1% Nonidet P-40. Cellular debris is

20   removed by centrifugation twice at 15,000 x g for 15 min. The cell lysate is incubated with 25 μl of M2 beads (Sigma) for 2 h at 4 °C with gentle rocking.

After extensive washing with lysis buffer, proteins bound to the beads are solubilized by boiling in SDS sample buffer, fractionated by SDS-polyacrylamide gel electrophoresis, transferred to polyvinylidene difluoride membrane and blotted with the indicated

25   antibodies. The reactive bands are visualized with horseradish peroxidase coupled to the appropriate secondary antibodies and the enhanced chemiluminescence (ECL) Western blotting detection system (Amersham Pharmacia Biotech).

V. Expression analysis

30   All cell lines used in the following experiments are NCI (National Cancer Institute) lines, and are available from ATCC (American Type Culture Collection, Manassas, VA 20110-2209). Normal and tumor tissues were obtained from Impath, UC Davis, Clontech, Stratagene, and Ambion.

TaqMan analysis was used to assess expression levels of the disclosed genes in various samples.

RNA was extracted from each tissue sample using Qiagen (Valencia, CA) RNeasy kits, following manufacturer's protocols, to a final concentration of 50ng/μl. Single stranded cDNA was then synthesized by reverse transcribing the RNA samples using random hexamers and 500ng of total RNA per reaction, following protocol 4304965 of Applied Biosystems (Foster City, CA, website at www.appliedbiosystems.com).

Primers for expression analysis using TaqMan assay (Applied Biosystems, Foster City, CA) were prepared according to the TaqMan protocols, and the following criteria: a) primer pairs were designed to span introns to eliminate genomic contamination, and b) each primer pair produced only one product.

Taqman reactions were carried out following manufacturer's protocols, in 25 μl total volume for 96-well plates and 10 μl total volume for 384-well plates, using 300nM primer and 250 nM probe, and approximately 25ng of cDNA. The standard curve for result analysis was prepared using a universal pool of human cDNA samples, which is a mixture of cDNAs from a wide variety of tissues so that the chance that a target will be present in appreciable amounts is good. The raw data were normalized using 18S rRNA (universally expressed in all tissues and cells).

For each expression analysis, tumor tissue samples were compared with matched normal tissues from the same patient. A gene was considered overexpressed in a tumor when the level of expression of the gene was 2 fold or higher in the tumor compared with its matched normal sample. In cases where normal tissue was not available, a universal pool of cDNA samples was used instead. In these cases, a gene was considered overexpressed in a tumor sample when the difference of expression levels between a tumor sample and the average of all normal samples from the same tissue type was greater than 2 times the standard deviation of all normal samples (i.e., Tumor – average(all normal samples) > 2 x STDEV(all normal samples)).

Results are shown in Table 1. Data presented in bold indicate that greater than 50% of tested tumor samples of the tissue type indicated in row 1 exhibited over expression of the gene listed in column 1, relative to normal samples. Underlined data indicates that between 25% to 49% of tested tumor samples exhibited over expression. A modulator identified by an assay described herein can be further validated for therapeutic effect by administration to a tumor in which the gene is overexpressed. A decrease in tumor growth confirms therapeutic utility of the modulator. Prior to treating a patient with

the modulator, the likelihood that the patient will respond to treatment can be diagnosed by obtaining a tumor sample from the patient, and assaying for expression of the gene targeted by the modulator. The expression data for the gene(s) can also be used as a diagnostic marker for disease progression. The assay can be performed by expression

5   analysis as described above, by antibody directed to the gene target, or by any other available detection method.

TABLE I

Gene Name	Breast Matched	Breast Total	Colon Matched	Lung Matched	Ovary Matched	Ovary Total	Kidney Matched
RRP1	0 / 3	0 / 5	11 / 26	8 / 14	1 / 4	3 / 14	2 / 19
RRP2	1 / 3	1 / 5	3 / 26	7 / 14	2 / 4	9 / 14	12 / 19
RRP3	0 / 3	0 / 5	4 / 26	4 / 14	1 / 4	3 / 14	4 / 19
RRP4	0 / 3	0 / 5	3 / 26	3 / 14	1 / 4	5 / 14	4 / 19
RRP5	0 / 3	0 / 5	6 / 26	1 / 14	1 / 4	1 / 14	4 / 19
RRP6	2 / 3	3 / 5	10 / 26	5 / 14	0 / 4	0 / 14	13 / 19
RRP7	0 / 3	0 / 5	6 / 26	1 / 14	1 / 4	1 / 14	1 / 19
RRP8	1 / 3	1 / 5	6 / 26	5 / 14	0 / 4	0 / 14	3 / 19

10

VI. Cell Biology and Functional Characterization

RRP Localization. In Drosophila, rhomboid is a cell surface protein. However, localization of human RRPs has not happened to date. Furthermore, in humans, it is not

15   known whether each of the termini of the rhomboids is intracellular or extracellular. To answer these questions, RRP1 was subcloned into pcDNA expression vector (Invitrogen) in frame with myc-his tag at the c-terminus, according to manufacturer's protocols. The resulting expression vector was transiently transfected using Lipofectamine Plus reagent (Life technologies) into human embryonic kidney HEK-293 cells. Immunofluorescence

20   staining using anti-myc antibody was then carried out on the cells to localize RRP1. Results of these experiments indicated that RRP1 is expressed at the cell surface. Furthermore, permeabilized (0.1% triton in PBS) and non-permeabilized cells show the same staining pattern, indicating that the c-terminus of the protein is outside, and the N-

terminus, inside. RRP2 – RRP8 are also subcloned and tagged at the c-terminus for localization of each protein, and also to assess the direction of the protein ends in each case.

5   **Involvement of RRP in EGFR signaling pathway.** While Drosophila rhomboid is essential in the EGFR pathway, the function of human RRPs have not been assessed to date. To assess the role of human RRPs in the EGFR signaling pathway, RRP1 was stably transfected using CaPO4 transfection kit (Clontech) into Hela cells, which have endogenous EGFR activity. Pooled stable cells were then examined in the following

10   assays.

EGFR activation as measured by tyrosine phosphorylation

Hela cells overexpressing RRP1 and parental Hela cells were tested for EGFR tyrosine phosphorylation by immunoblot against anti-phosphotyrosine using anti-phosphotyrosine antibody (Upstate Biotechnology). Cells overexpressing RRP1

15   demonstrated up to two-fold increase in EGFR tyrosine phosphorylation as compared to parental Hela controls.

EGFR expression in cells overexpressing RRP1.

Expression of EGFR was increased when tested by immunoblot using anti-EGFR antibody against direct cell lysates in cells overexpressing RRP1. Furthermore, expression

20   of EGFR was also increased in tumor samples overexpressing RRP1, by TaqMan® analysis. The correlation of both in-vitro and in-vivo results for concordant overexpression of EGFR and RRP1 is quite significant, and provides a tool for assessing the functional relationship of the two proteins in any tumor sample or cell line.

Cell proliferation and Migration.

25   Hela cells overexpressing RRP1 demonstrate a two-fold increase in cell proliferation compared to Hela parental cells as measured by Procheck (Intergen) and AlamarBlue (Biosource International) chemiluminescent assays. In addition, Fluroblok (BD Biosource) analysis showed a slight increase in motility in cells overexpressing RRP1 compared to Hela parental cells.

30   Taken together, these results demonstrate a strong functional conservation of rhomboids across evolution. Furthermore, increased RRP1 expression in tumor samples and cell lines, concurrent increased expression of EGFR, and increased proliferation and motility of cells expressing RRP1 suggest involvement of RRP1 in various cancers. RRP1 expression levels can thus be used to screen for tumors with defective EGFR pathways, to

identify tumors amenable to treatment with the compounds and antibodies identified by the methods of the present invention. Same aforementioned experiments are performed with RRP2 – RRP8 to assess their involvement in the EGFR signaling pathway.

5 **RRP1 binding target.** In Drosophila, rhomboid transforms Spitz (homolog of human TGF-α) from a membrane-bound to a secreted form, and thus triggers and upregulates the DER (Drosophila EGFR) pathway. No such function has been shown for human rhomboids. HEK293 cells were transiently co-transfected with RRP1 and TM-TGF- α transmembrane-bound TGF), then immunoprecipitated with anti-myc antibody
10 (for RRP1) and immunoblotted with anti-TGF-α. Cells transfected with vector alone, RRP1 alone, or TM-TGF-α alone were used as controls. A band corresponding to the size of TGF- α was recognized by anti-TGF- α antibody, while no bands were detected in the control lanes. This data indicates that RRP1 physically interacts with TGF- α. The same experiments are performed with RRP2 – RRP8 to determine their binding partners.
15

Effect of RRP1 on HB_EGF and ERK phosphorylation. RRP1 was transfected into HEK293 embryonic kidney cells to investigate effects of RRP1 on HB-EGF (an EGFR ligand) and ERK phosphorylation. The RRP1 full-length clone was inserted into the mammalian expression vector pCDNA3.1 (Invitrogen) with N-terminal FLAG,
20 according to manufacturer's protocol. HEK293T cells were seeded into 6-well plates at 600,000 cells per well, grown in DMEM with 10% Fetal Bovine Serum (FBS) and a Penicillin/Streptomycin cocktail. Seeded cells were allowed to grow for 24 hours at 37 °C and 5% CO_2. Cells were transfected with 2 μg of RRP1 plasmid per well using LipofectAmine PLUS reagent (Life Technologies) according to manufacturer's protocol.
25 Three hours after transfection, FBS concentration was brought up to 10%,total volume was increased to 4 ml with DMEM, and cells were placed back into incubator for 48 hours. Transfected cells were then processed for Western analysis. Cells were lysed in NP40 lysis buffer containing protease and phosphatase inhibitors. Protein concentrations were determined with Pierce BCA Protein Assay Reagent, and 20 μg of protein was added
30 to each lane on a 10% NUPAGE Bis-Tris Gel (Invitrogen), run and transferred to nitrocellulose paper, following instructions in the NUPAGE Bis-Tris instruction booklet. To block, blots were soaked in 10ml of 0.5% non-fat milk for 2 hours, and washed. Primary antibodies were then added at appropriate dilution (diluted in binding buffer): Mouse α-FLAG: 1:1000; Mouse α-HBEGF: 1:1000 (R&D Systems); Rabbit α-Total

ERK: 1:1000 (Cell Signaling); Rabbit α-Phosphorylated ERK: 1:1000 (Cell Signaling).
Blots were rinsed and secondary antibodies were added: Donkey α-rabbit: 1:10000
(Jackson ImmunoResearch, HRP conjugated); Donkey α-mouse: 1:10000 (Jackson
ImmunoResearch, HRP conjugated). Blots were then rinsed, incubated 1:1 Supersignal

5   West Pico Stable Peroxide Solution and Luminol /Enhancer solution (Pierce) for 5
minutes according to instructions, and exposed to film. In these experiments, RRP-1
transfected HEK-293T cells showed elevated levels of HB-EGF, and elevated ERK
phosphorylation.

In Drosophila, RRP1 has been shown to increase EGFR signaling through cleaving

10   and releasing activated EGFR ligand (spitz; described *supra*). Our findings are consistent
with those reports. Transient transfection of human RRP1 in HEK293T cells clearly
induced ERK phosphorylation, a signaling event down-stream of EGFR activation. In
addition, co-transfection of RRP1 and HB-EGF (one of EGFR ligands in human) in
HEK293T cells yielded higher HB-EGF protein level than HB-EGF transfection along,

15   suggesting RRP1 may increase the overall level of EGFR ligands in cells. Overall, these
experiments suggest that human RRP1 increases EGFR signaling.

Treatment of RRP1transfected cells with EGFR and HB-EGF inhibitors. The
aim of this experiment was to test the effect of EGFR and HB-EGF inhibitors on

20   phosphorylated ERK levels in RRP1 transfected cells. HEK293T cells were grown and
transfected as described above. At 24 and 44 hours after transfection, the following EGFR
or HB-EGF inhibitors were added to wells: for HB_EGF: HB-EGF mouse monoclonal
antibody (R&D Sysytems; 10ug/ml), and CRM197 (10ug/ml). For EGFR: Ab225 mouse
monoclonal antibody (ATCC; 5ug/ml), and 168393 small molecule inhibitor (Calbiochem;

25   2.7uM). At 48 hours post transfection, cell lysates were collected and analyzed by western
analysis, as described above. Both EGFR inhibitors (Ab225 and 168393) decreased the
levels of pERK in RRP1 transfected cells. Anti HBEGF antibody also was able to reduce
pERK levels in RRP1 transfected cells. The HBEGF inhibitor CRM197 did not have an
affect. This result may be potentially useful in designing an assay to test

30   antibodies/compounds that affect the MAPK signaling cascade. Based on the observed
results, we propose the following hypothesis: human RRP1 increases EGFR/Erk signaling
through cleavage/shedding of HB-EGF. EGFR signaling is a pivotal factor in
tumorigenesis, and HB-EGF is one of the most important growth factors that drive tumor

development through autocrine and other mechanisms. Therefore, RRP1 may play an important role in cancer and other diseases that link to irregulation of EGFR signaling.

Antibodies against human rhomboids. A peptide antibody, SEQ ID NO: 17 , and referred to as peptide 1, was made against the N-terminus of RRP1. This antibody

5   displayed strong affinity for denatured RRP1, such as in cells fixed with formalin or proteins on western blots. In addition, staining pattern remained the same when permeabilized cells were co-stained with anti-myc antibodies (for the c-terminus of RRP1) and peptide 1, suggesting that Peptide 1 does recognize RRP1. Furthermore, the cells were significantly immunostained only when they were permeabilized, suggesting that the

10   N-terminus of RRP1 is inside the cells. This result correlates with the previous staining pattern using anti-myc antibody for the C-terminus.

RRP1 protein expression in tumors was examined by Immunohistochemistry (IHC) using peptide 1. Positive staining on epithelial cells of colon adenocarcinoma tissue sections were observed as compared to background staining in preimmune serum,

15   suggesting the antibody is specific for its target, RRP1, in colon cancer cells. Moreover, increased staining of colorectal adenocarcinomas was observed compared to matched normal tissues using peptide 1 antibody. Taken together, these results suggest that peptide 1 antibody is a powerful tool to assess expression levels of RRP1 in tumor samples. Antibodies are also produced against the rest of RRPs to assess their expression levels in

20   tumors.

VII. Mouse RRP1 (mRRP1) gene structure

A BAC shotgun sequencing approach (for a review, see Green ED. (2001) Nature Review Genetics 2:573-583) was used to determine the intron-exon structure of the

25   murine RRP1 gene. Based upon the genebank entries of a partial rat cDNA and the human gene structure, a PCR primer pair was designed that allowed the amplification of a part of the murine RRP1 gene from mouse genomic DNA. (RRP1s (SEQ ID NO:48): 5′-CACCCACCAGCCGCACTGGTC; RRP8as (SEQ ID NO:49): 5′-GAGCAGCTAGGGTTCAATG, 95°C 120sec, 95°C 45sec, 60°C 45sec, 72°C 45sec,

30   40cycles, 1.5mM MgCl2). This PCR primer set was then used to screen a commercially available ES-129SvJ BAC library from Incyte Genomics (St. Louis, USA). 2 PCR positive clones were identified (clone addresses 178/O15 and 195/D15). Clone 195/D15 was chosen to construct a shotgun sequencing library according to standard procedures (TOPO Shotgun Subcloning Kit, #K-7000-01, Invitrogen). Upon sequence assembly, a

contiguous 30 Kilobases (kb) genomic sequence was identified that carried the entire mRRP1 gene. Subsequently, the DNA from clone 195/D15 BAC served as a PCR template to amplify mRRP1 genomic regions used for the gene targeting approach. The mRRP1 gene contains 8 exons, which span a genomic region of at least 2.6 kb. The

5   translation start on the sequence stretch defined as exon 1 has not been identified so that additional 5′ exons may exist.

VIII. Generation of the mice

Targeting Strategy

10   We utilized a targeting strategy that allows sequential deletion of vector sequences in the same cells. Chimeras generated by injection of mRRP1 targeted ES-cells can be used to create i) KO mice harboring a completely inactivated mRRP gene by breeding to Cre-deleter mice (described below) as well as ii) mice with functional mRRP1 that can be deleted in a tissue specific or time specific manner by breeding to FLP-deleter mice

15   (described below).

This strategy has the advantage that the NEO selection cassette is removed in both chimeras, thus avoiding potential hypomorphic effect of the selection marker.

Targeting vector

20   The targeting vector, comprising the nucleotide sequence to be incorporated into the wild –type (WT) genomic sequence, and one or more selectable markers (FIG. 1) was made based on a BAC genomic clone from the mouse strain 129SvJ background obtained from Incyte Genomics. The mRRP1 targeting vector was designed such that exons 2-8 of the mRRP1 gene were flanked with Cre recombinase recognition sequences (loxP sites)

25   after homologous recombination. Exons 2-8 of the murine RRP encode the rhomboid domain which provide protein activity, therefore a deletion of exons 2-8 results in a null allele. Three genomic fragments were amplified by PCR and cloned individually in reversed orientation into the base vector pEasyfloxFRTGK12 (FIG. 1). This base vector has: an ampicillin resistance gene; an orign of replication; a pgtN29 plasmid (New

30   England Biolabs), used as the source of pgk neo; pGT60mcs plasmid (InvivoGen), used as the source for TK reading frame; and LoxP and FRT sites, introduced by synthetic oligonucleotides. The 5′ region of homology (fragment A) amplified with oligo 164 (SEQ ID NO:62) and oligo 144 (SEQ ID NO:50) covered a 3.5 kb fragment that was inserted into PacI / PmeI sites upstream the 2nd loxP site of pEasyfloxFRTGK12. A 2.5 kb 3′

fragment encompassing exons 2-8 (fragment B) and amplified with oligos 145 (SEQ ID NO:51) and 146 (SEQ ID NO:52) was cloned into a AscI / FseI I sites downstream of 2nd loxP site and upstream of 2nd FRT site and *PGK-neo* cassette of pEasyfloxFRTGK12. An additional 1.0 kb fragment further 3′ (fragment C) and amplified with oligos 147 (SEQ ID

5   NO:53) and 148 (SEQ ID NO:54) was inserted into NotI/SgfI sites located downstream the 1st FRT and 1st loxP site of pEasyfloxFRTGK12. The loxP- and FRT-flanked neomycin resistance cassette downstream of exons 2-8 of the RRP gene allowed genomic selection for homologous recombination and a *PGK-thymidine kinase* gene cassette (PGK-TK) inserted at the 5′ end of the construct was used to select against random integration

10   events.

In order to establish a positive control for the PCR screen on ES cell clones showing homologous recombination events, an additional 1.25 kb fragment was amplified using oligo 147 (SEQ ID NO:53) and oligo 149 (SEQ ID NO:55) and cloned into NotI / SgfI sites in pEasyfloxFRTGK12. This fragment extended the 3′ homology region of the

15   final targeting vector by 250 base pairs, allowing establishment of PCR screening with a vector specific oligo 174 (SEQ ID NO:64) and RRP gene specific oligo 169 (SEQ ID NO:63) located 3' external in relation to the targeting vector.

Generation of targeted ES-cells

20   Three parental embryonic stem (ES) cell lines **B6-2** (EUROGENETEC, Ref: TG-ES01-02, derived from mouse strain C57BL/6 N TacfBr), **SvEv** (EUROGENETEC, Ref: TG-ES01-01, derived from mouse strain 129S6/SvEvTacfBr), and **E-14** ((Hooper et al., (1987) *Nature (London)* 326, 292-295); derived from mouse strain 129P2/OlaHsd) were employed for targeting of the mRRP1 locus.

25   All ES cell lines were cultured and electroporated under the conditions described in Torres, R. M. & Kuehn, R. Laboratory Protocols for Conditional Gene Targeting, *Oxford, New York, Tokyo:* 1997. In brief, for electroporation 1 x 10^7 cells were mixed with 30 μg of *SfiI* linearized targeting DNA in a total volume of 800 μl transfection buffer and electroporated using a Bio-Rad Gene Pulser (0,24 kV, 500 μF, 0,4 cm electrode

30   distance). Cells were reseeded on 10-cm tissue culture dishes, containing selection resistant mouse embryonic fibroblast feeder cells (prepared from targeted mouse strain IL4 pepneo; Mueller et al., (1992) *Eur. J. Immunol.* 21: 921 –25) at a density of 2.5 x 10^6 cells per plate. Drug selection (G418 (Gibco/BRL # 10131-019)), 200 mg/ml) was started after 1 day. ES cell clones were negatively selected the 3rd day after electroporation with the

antibiotic Gancyclovir (Cymeven, Roche), 2 μM). After 8 days, resistant colonies were randomly picked and individually expanded. Correctly targeted ES clones were obtained for all 3 ES cell lines employed in the experiment.

5   Successful targeting of the mRRP gene was detected in resistant ES clones by PCR-screening with primer oligo 174 (SEQ ID NO:64) and oligo 169 (SEQ ID NO:63). Corecombination of the distal loxP site was detected by using oligo 200 (SEQ ID NO:65) and oligo 201 (SEQ ID NO:66). PCR-positive clones were expanded and confirmed by EcoRI / SfiI restriction enzyme digestion and Southern hybridization employing the 5' external probe amplified with oligos 152 (SEQ ID NO:56) and 153 (SEQ ID NO:57).

10   Upon homologous recombination, the wild type 10.0 kb band was reduced to 7.5 kb (Figure 2). Corecombination of the 2nd loxP site in homologous recombinants was confirmed by HindIII and 3' external probe amplified with oligos 160 (SEQ ID NO:60) and 162 (SEQ ID NO:61)(Figure 2). Upon corecombination of the 2nd loxP site, the wildtype 4.3 kb band was increased to 5.8 kb (Figure 2). Single integration of the

15   targeting vector was confirmed by HindIII digestion and Southern hybridization using the internal probe amplified with oligos 154 (SEQ ID NO:58) and 155 (SEQ ID NO:59)(Figure 2).

Table 2 provides a summary of these experiments.

20   Table 2:Homologous recombination of the RRP-1 targeting construct in various ES cells

ES Cell	strain	# cells x 10^7	G418 res. anal. Clones		HR	2. loxP #	Clone
B6-2	C57BL/6	4	2000	876	3	2	G-D5 E-D9
SvEv	129S6	2	Not counted	95	1	0	
		3	Not counted	295	1	1	D-E7
E-14	129P2	6	Not counted	641	11	3	B-E2 E-F2 C-C8

Generation of mRRP1 chimeras

Chimeras were generated by injection of targeted ES cells into Balb/C host embryos (Hogan, B., Beddington, R., Costantini, F. & Lacy, E. eds. Manipulating the Mouse Embryo, a Laboratory Manual. 2nd ed. Cold Spring Harbor, New York: Cold

5   Spring Harbor, Laboratory Press, 1994). Germline transmission was obtained by breeding chimeric mice to C57BL/6 females, resulting in mRRP1 heterozygous animals on a pure C57BL/6 strain background and hybrid (C57B/L6 x 129SvEv)$_{F1}$ and (C57B/L6 x 129P2/Ola)$_{F1}$ background.

Resulting animals were typed by Southern blot hybridization on tail DNA after

10   digestion with HindIII with the 3′ external probe. In heterozygote animals for the target allele, the 4.3kb wild type band was increased to 5.8 kb.

Table 3 provides a summary of these experiments.

Table 3: Summary of chimera production and germline transmission of the target allele.

15

Clone	blastocysts transf.	pups born	# of chimeras	GL transmis.	heterozyg. mice
B6-2					
G-D5	246	37	32	10/11	yes
E-D9	372	103	70	6/10	yes
Total	618	140	102	17/21	
SvEv					
D-E7	302	71	40	8/10	yes
E-14					
B-E2	64	17	14	8/10	yes
Grand Total	984	228	156	33 / 41	

Heterozygous animals by in vivo deletion.

Target allele with Cre directed deletion

An in vivo deletion approach was pursued to generate C57BL/6 mice carrying the

20   target alleles with Cre-directed deletion from mRRP1. Using this approach, the Cre transgene directs the removal of the sequences flanked by the LoxP sites, and thus, creates an RRP1 knockout in every cell of the animal (FIG.3).

Heterozygote mice carrying the mRRP1 target allele (mRRP1^{targ/+}) from the B6-2 clones G-D5 and E-D9 (Table 2) were bred with transgenic C57BL/6-Cre-Deleter mice

25   (Schwenk et al, NAR (95) Vol. 23, No 24, 5080-5081). This deleter mice have been crossed back to B6 for 10 generations, to establish the transgene on an congenic C57BL/6

48

background. This was done to maintain the C57BL/6 background after the in vivo deletion of the mRRP1 targeted allele. Resulting pups were genotyped on tail DNA for the presence of the target allele with Cre-directed deletion of mRRP1 by Southern Blot hybridization and by PCR for the presence of the Cre transgene.

5   Tail DNA was digested with HindIII and hybridized with the 3´ external probe. In heterozygote animals the 5.8 kb band was reduced to a 1.3 kb band for the target allele. Alternatively animals were typed by PCR.

Target allele with Flp-directed deletion

10   An in vivo deletion approach was pursued, to generate C57BL/6 mice carrying the target allele with Flp directed deletion of the Neo marker from RRP1. In this case, the mRRP1 gene is still functional in every cell, but the animals may later be crossed with conditional Cre deleter mice for tissue specific or time-specific deletion of the sequences flanked by LoxP to create conditional mRRP1 knockout mice.

15   To avoid any potential hypomorphic effects, the Frt flanked Neo selection marker was removed in vivo by Flp recombination. Heterozygote mice carrying the mRRP1 target allele from the B6-2 clones G-D5 and E-D9 (Table 2) were bred with transgenic C57BL/6 mice heterozygous for the CAAGS-FLPe transgene on a pure B6 background (for FLPe description, see Rodriguez et al., (2000) Nat Genet 25:139-140). Resulting
20   pups were genotyped on tail DNA for the presence of the target allele by Southern Blot hybridization and by PCR for the presence of the CAAGS-Flp transgene.

           Tail DNA was digested with HindIII and hybridized the 3´ external probe. In heterozygote animals the 4.3 kb target allele band was reduced to a 3.8 kb band after the removal of the Frt-flanked Neo marker (FIG.4). Alternatively animals were typed by
25   PCR.

Homozygous animals
Target allele with Cre directed deletion

           To generate mice homozygous for the mRRP1 knockout allele on the C57BL/6
30   background, heterozygous animals derived from the in vivo deleted clones G-D5 and E-D9 were intercrossed. From this cross 17 litters were born, giving rise to 106 pups. 28 of these pups died pre-weaning. Two of these pups were recovered and genotyped by PCR as homozygous complete KO. 49 of the living animals were weaned and 14 animals were genotyped on tail DNA with the 3´ external probe after digestion with HindIII. In

homozygous animals the original 4.3 kb wild type band was reduced to a 1.3 kb band. Alternatively animals were typed by PCR.

From the 14 animals 3 typed as homozygous mutant, 7 typed as heterozygous mutant and 4 typed as wildtype. These numbers matched the expected Mendelian

5   distribution. The homozygous animals appear to be normal. No gross difference was detected in comparison to their wildtype littermates.

To test for the fertility of the homozygous mutant animals and to increase further the colony, homozygous mutant animals for the mRRP1 derived from the in vivo deleted clones G-D5 were intercrossed. To date, 14 pups have been born from these intercrosses,

10   and 4 have been weaned, demonstrating the fertility of the homozygous mRRP1 KO mice.

The homozygous mRRP1 KO mice and cells are used to analyze mRRP1 function in mammals. Data from Example VI show that overexpression of human RRP1 in HeLa cells results in enhanced phosphorylation of EGFR. Inactivation of mRRP1 gene function should therefore result in attenuated phosphorylation of EGFR and or MAPK in tissues

15   obtained from mRRP1 KO mice.

Growth factors like EGF, PDGF cause proliferation in fibroblast cell culture. As Drosophila Rhomboid and mammalian RRPs are thought to release membrane-bound growth factors by cleavage, RRP function with regard to cell cycle progression and proliferation is analyzed with embryonic and adult fibroblasts isolated from mRRP1 KO

20   mice.

Direct effects of RRP function on tumor development is analyzed by crossing mRRP1 KO mice with APCMin mice, a model for human colon cancer. As the mRRP1 KO strain and the APCmin (Moser AR et al., (1993) Proc Natl Acad Sci U S A 90:8977-81) strain are both on C57BL/6 background, tumor formation can be analyzed on pure

25   genetic backgrounds, and thus side effects resulting from mixed genetic backgrounds can be excluded. A strong reduction of small intestinal tumor formation in mRRP1 -/-, APCMin +/- mice versus APCMin +/- mice is expected, demonstrating that mRRP1 is a powerful target for anti-cancer drug development.

RRP is located at the cell surface and therefore a target for antibody therapy. As

30   the extracellular loops of mouse and human RRP protein are strongly conserved, it is difficult to generate antibodies against human RRP in mice, as mice will recognize the human RRP as self. Therefore, RRP KO mice are used for the production of monoclonal antibodies directed against human RRP protein.

35   50

WHAT IS CLAIMED IS:

1. A method of specifically inhibiting growth of tumor cells that overexpress an RRP protein comprising contacting said tumor cells with an amount of an exogenous RRP

5   binding agent that results in inhibition in growth of tumor cells.

2. The method of claim 1 wherein said tumor cells are from a tissue selected from the group consisting of breast, ovary, lung, kidney, and colon.

10   3. The method of claim 1 wherein said RRP binding agent is an antibody or a fragment of said antibody.

4. The method according to claim 3 wherein said antibody is humanized.

15   5. The method according to claim 3 wherein said antibody comprises human constant region residues and rhomboid-specific mouse variable region residues.

6. The method according to claim 3 wherein said antibody is bound to a detectable label.

20   7. The method according to claim 3 wherein said antibody is bound to a cytotoxic label.

8. A method of screening for agents that modulate the interaction of an RRP polypeptide with an RRP binding target, comprising:
      a) expressing a recombinant RRP polypeptide,
25   b) incubating the recombinant RRP polypeptide with an RRP binding target and a candidate RRP modulating agent and
      c) determining whether said candidate RRP modulating agent modulates the binding of the RRP polypeptide with the RRP binding target.

30   9. The method according to claim 8 wherein said binding target is selected from the group consisting of TGFα, EGF, and amphiregulin.

10. The method according to claim 8 wherein said binding target is TGFα.

51

11. The method according to claim 8 wherein said candidate RRP modulating agent is an antibody.

12. The method according to claim 8 wherein said candidate RRP modulating agent is a
5   small organic molecule.

13. A method of diagnosing a tumor cell as having abnormal p53 or p21 pathway signaling, comprising measuring RRP expression levels in said tumor cell, wherein increased RRP expression levels in the tumor cell compared to a non-tumor cell of the
10   same tissue or organ as the tumor cell, indicates defective p53 or p21 pathway signaling in the tumor cell.

14. A method of identifying a candidate p53 or p21 pathway modulating agent, said method comprising the steps of:
15   a) providing an assay system comprising a purified RRP polypeptide or nucleic acid or a functionally active fragment or derivative thereof;
         b) contacting the assay system with a test agent under conditions whereby, but for the presence of the test agent, the system provides a reference activity; and
         c) detecting a test agent-biased activity of the assay system, wherein a difference
20   between the test agent-biased activity and the reference activity identifies the test agent as a candidate p53 or p21 pathway modulating agent.

15. The method of Claim 14 wherein the assay system comprises cultured cells that express the RRP polypeptide.
25

16. The method of Claim 15 wherein the cultured cells additionally have defective p53 or p21 function.

17. The method of Claim 14 wherein the assay system includes a screening assay
30   comprising a RRP polypeptide, and the candidate test agent is a small molecule modulator.

18. The method of Claim 17 wherein the assay is a protease assay.

52

19. The method of Claim 14 wherein the assay system is selected from the group consisting of an apoptosis assay system, a cell proliferation assay system, an angiogenesis assay system, and a hypoxic induction assay system.

5   20. The method of Claim 14 wherein the assay system includes a binding assay comprising a RRP polypeptide and the candidate test agent is an antibody.

21. The method of Claim 14 wherein the assay system includes an expression assay comprising a RRP nucleic acid and the candidate test agent is a nucleic acid modulator.

10

22. The method of claim 21 wherein the nucleic acid modulator is an antisense oligomer.

23. The method of Claim 21 wherein the nucleic acid modulator is a PMO.

15   24. The method of Claim 14 additionally comprising:
           d) administering the candidate p53 or p21 pathway modulating agent identified in (c)
        to a model system comprising cells defective in p53 or p21 function and, detecting a
        phenotypic change in the model system that indicates that the p53 or p21 function is
        restored.

20

25. The method of Claim 24 wherein the model system is a mouse model with defective p53 or p21 function.

26. A method for modulating a p53 or p21 pathway of a cell comprising contacting a cell
25   defective in p53 or p21 function with a candidate modulator that specifically binds to an
        RRP polypeptide comprising an amino acid sequence selected from group consisting of
        SEQ ID NOs:2, 4, 6, 8, 10, 12, 14, 16, and 46 whereby p53 or p21 function is restored.

27. The method of claim 26 wherein the candidate modulator is administered to a
30   vertebrate animal predetermined to have a disease or disorder resulting from a defect in
        p53 or p21 function.

28. The method of Claim 26 wherein the candidate modulator is selected from the group consisting of an antibody and a small molecule.

29. The method of Claim 14, comprising the additional steps of:

 e) providing a secondary assay system comprising cultured cells or a non-human animal expressing RRP ,

 f) contacting the secondary assay system with the test agent of (b) or an agent derived therefrom under conditions whereby, but for the presence of the test agent or agent derived therefrom, the system provides a reference activity; and

 g) detecting an agent-biased activity of the second assay system,

 h) wherein a difference between the agent-biased activity and the reference activity of the second assay system confirms the test agent or agent derived therefrom as a candidate p53 or p21 pathway modulating agent,

 i) and wherein the second assay detects an agent-biased change in the p53 or p21 pathway.

30. The method of Claim 29 wherein the secondary assay system comprises cultured cells.

31. The method of Claim 29 wherein the secondary assay system comprises a non-human animal.

32. The method of Claim 31 wherein the non-human animal mis-expresses a p53 or p21 pathway gene.

33. A method of modulating p53 or p21 pathway in a mammalian cell comprising contacting the cell with an agent that specifically binds a RRP polypeptide or nucleic acid.

34. The method of Claim 33 wherein the agent is administered to a mammalian animal predetermined to have a pathology associated with the p53 or p21 pathway.

35. The method of Claim 33 wherein the agent is a small molecule modulator, a nucleic acid modulator, or an antibody.

36. A method for diagnosing a disease in a patient comprising:

 a) obtaining a biological sample from the patient;

 b) contacting the sample with a probe for RRP expression;

    c) comparing results from step (b) with a control;

    d) determining whether step (c) indicates a likelihood of disease.

37. The method of claim 36 wherein said disease is cancer.

38. The method of claim 37, wherein said cancer is a cancer as shown in Table 1 as having >25% expression level.

39. A purified nucleic acid molecule that encodes a polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:46, or reverse complement thereof.

40. The nucleic acid molecule of Claim 39 which is capable of hybridizing to a nucleic acid sequence of SEQ ID NO:45 using high stringency hybridization conditions.

41. A recombinant expression system comprising a DNA or RNA molecule, wherein said expression system is capable of producing an mRRP1 polypeptide comprising the amino acid sequence of SEQ ID NO:46 when said expression system is present in a compatible host cell.

42. A host cell comprising the expression system of claim 41.

43. A process for producing an mRRP1 protein comprising culturing the host cell of Claim 42 under conditions suitable for expression of said mRRP1 protein and recovering said protein.

44. A process for producing a cell which produces an mRRP1 protein comprising the transformation or transfection of a host cell with the expression system of claim 41 such that the host cell, under appropriate culture conditions, produces an mRRP1 protein.

45. A recombinant host cell expressing the protein produced by the method of claim 44.

46. A transgenic mouse whose genome comprises a disruption in an endogenous RRP gene wherein said disruption results in decreased expression or a lack of expression of said endogenous RRP gene.

47. The mouse of claim 46, wherein said mouse is homozygous for the disrupted RRP gene.

5   48. A cell isolated from the transgenic mouse of claim 46.

49. A mouse cell whose genome comprises a disruption in an endogenous RRP gene wherein said disruption results in decreased expression or a lack of expression of said endogenous RRP gene.

10

50. A method of selecting an agent that modulates cell proliferation comprising:

   a) providing a first and a second isolated mouse cell wherein the genome of both the first and second isolated mouse cell has been manipulated to comprise a disruption in an endogenous RRP gene wherein said disruption results in decreased expression or a

15   lack of expression of said endogenous RRP gene;

   b) administering an agent to said first isolated mouse cell ; and

   c) determining the amount of proliferation of the first and second cell, wherein a difference in the amount of proliferation of the first cell as compared to the second cell indicates that the agent modulates cell proliferation.

20

51. A method of making an antibody against a human RRP comprising:

   a) injecting RRP or an antigenic fragment thereof to mouse of claim 46; and

   b) recovering said antibody.

25

FIG.1 pEasyfloxFRTGK12 mRRP1

Sfi I (14298)

Not I (14290)

Fragment C

Sgf I (13285)

loxP

FRT

pgk NEO

FRT

Asc I (11243)

14304 bp

*Srf*I (2887)

pgkTK

Fragment B

I-SceI (5106)

Fse I (8719)

Pme I (5122)

loxP

Pac I (8640)

Fragment A
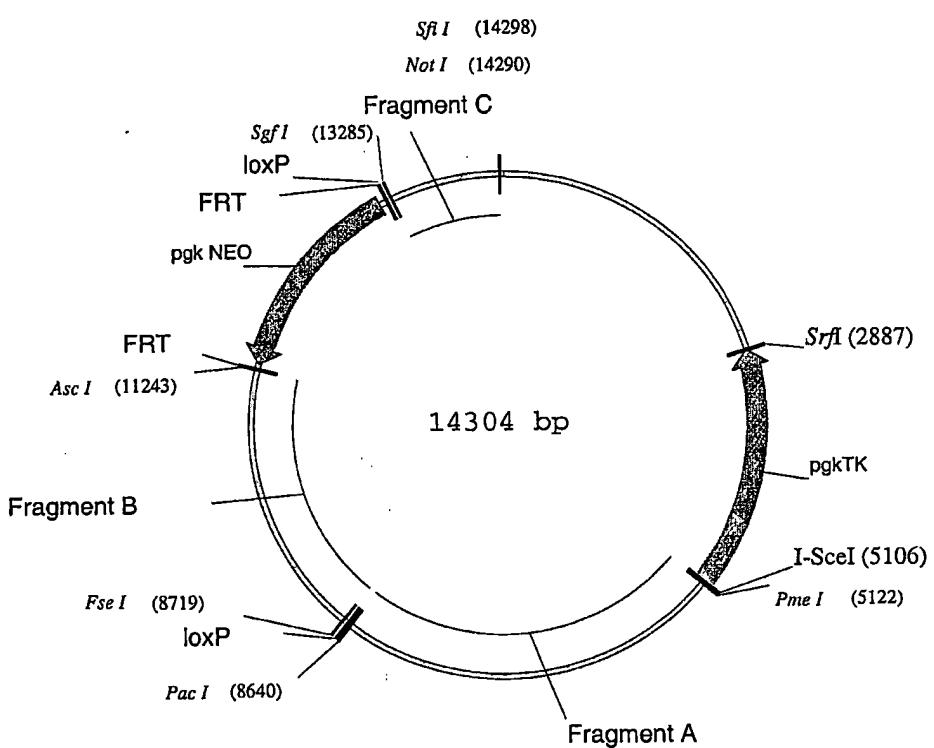
FIG.2 mRRP1 gene targeting
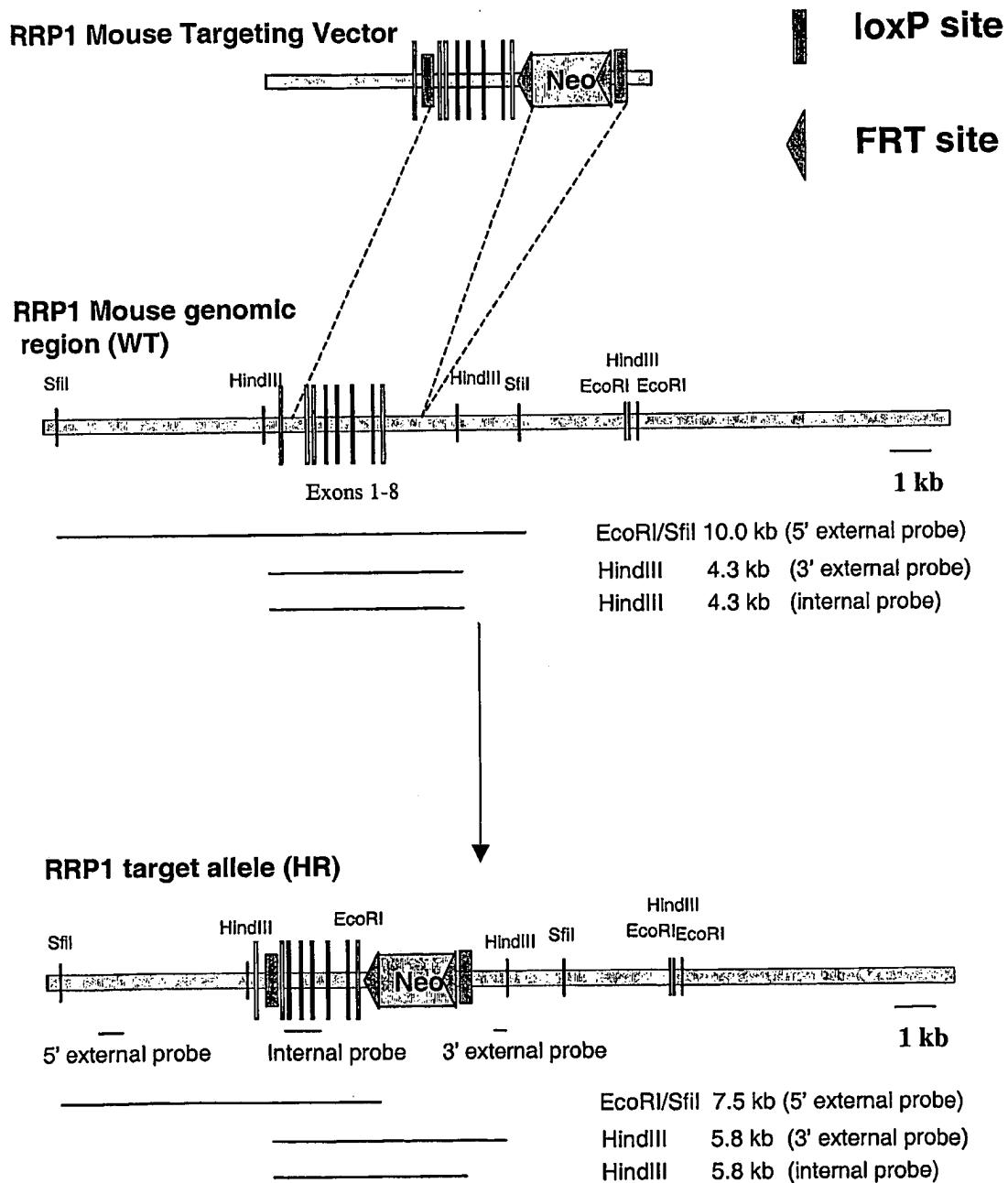
FIG.3
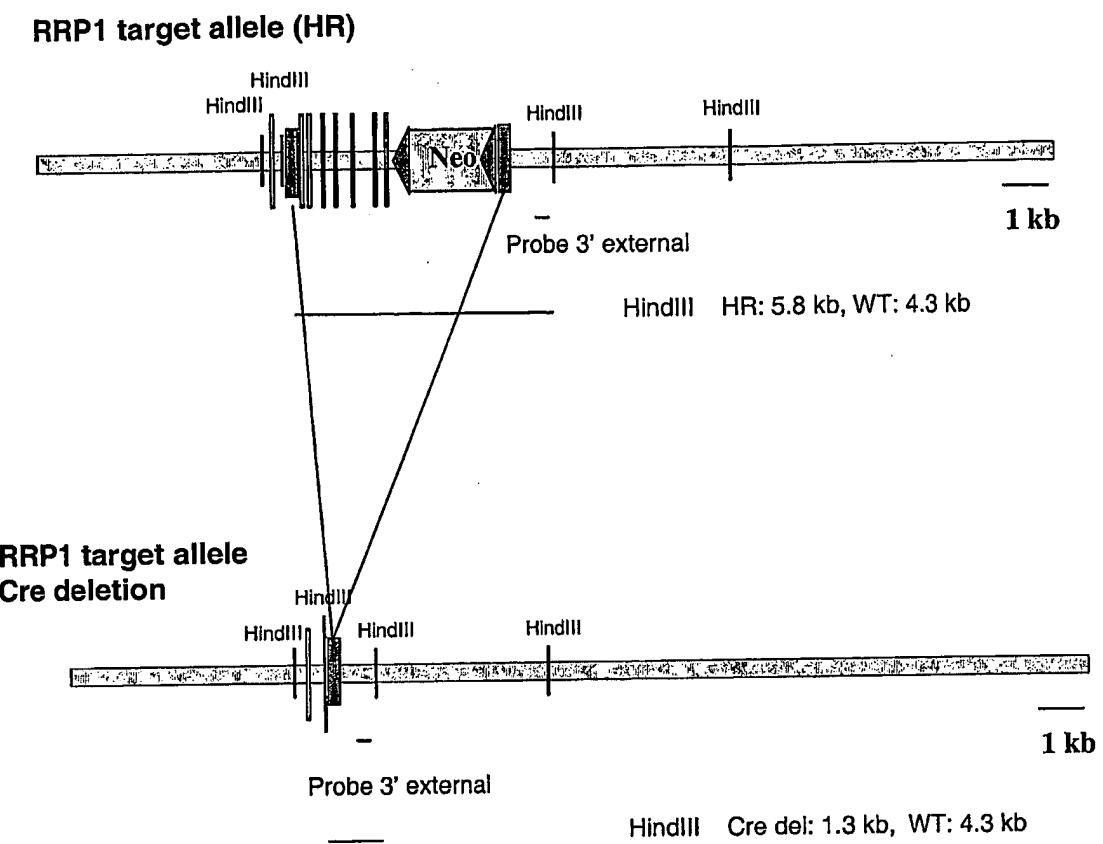
RRP1 target allele (HR)

Probe 3' external

HindIII   HR: 5.8 kb, WT: 4.3 kb
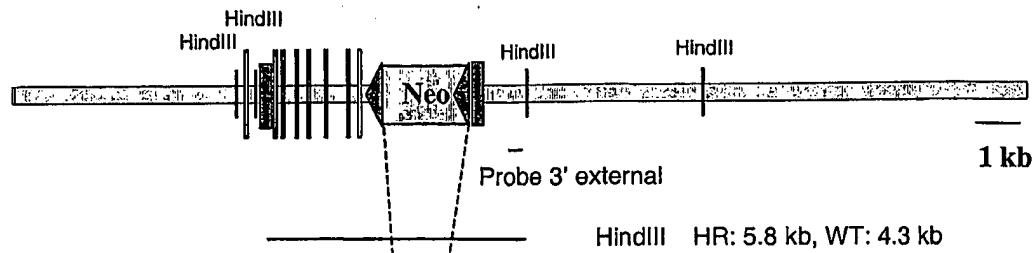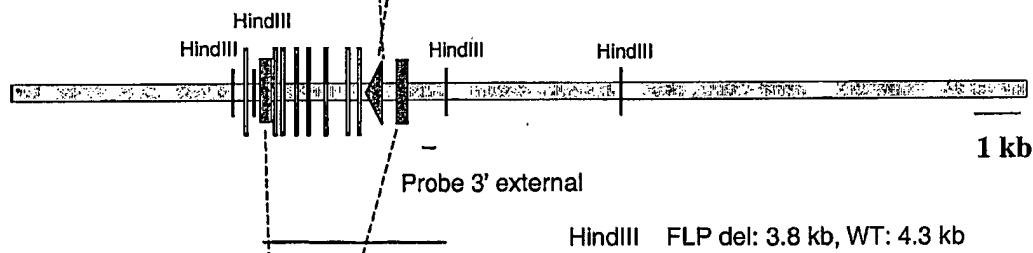
RRP1 target allele
Cre deletion

Probe 3' external

HindIII   Cre del: 1.3 kb, WT: 4.3 kb

FIG.4
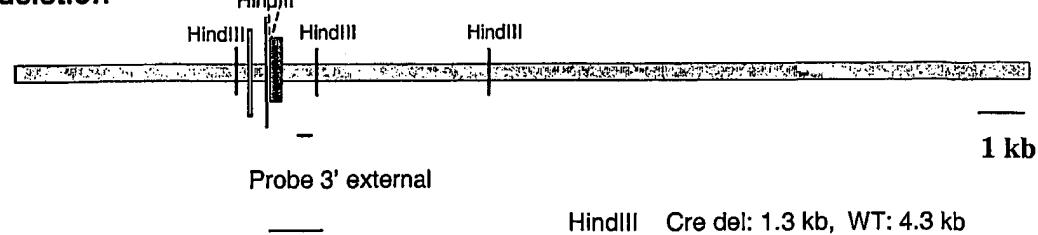
SEQUENCE LISTING

<110>   EXELIXIS, INC.

<120>   RRP SEQUENCES AND KNOCKOUT MICE AND USES THEREOF

<130>   EX02-004C-PC

<150>   US 10/056,790
<151>   2002-01-23

<150>   US 09/908,419
<151>   2001-07-18

<150>   US 60/219,289
<151>   2000-07-19

<150>   US 60/277,487
<151>   2001-03-21

<150>   US 60/304,863
<151>   2001-07-12

<150>   US 60/277,471
<151>   2001-03-21

<150>   US60/305,017
<151>   2001-07-12

<150>   US 60/296,076
<151>   2001-06-05

<150>   US 60/328,605
<151>   2001-10-10

<150>   US 60/328,491
<151>   2001-10-10

<160>   70

<170>   PatentIn version 3.2

<210>   1
<211>   1559
<212>   DNA
<213>   Homo sapiens

<400>   1
cttggacctt ggccctcgct ttccaggatg ggtagggtgg aagacggggg aacaactgag   60

gagctggagg actgggaccc aggcaccagt gccctgccag ctcctgggat caagcagggt   120

cccagggaac agacaggcac ggggcccctg tcccaaaagt gctgggagcc tgagcctgat   180

gctcccagcc agcctggccc agccctttgg tccaggggtc gggcccgcac tcaggccttg   240

gctggcggct cctcactgca gcagctggac cccgagaaca caggcttcat cggtgcggac   300

accttcactg gcctggtgca cagccatgag ctgcccctgg acccggccaa gctggacatg   360

ctggtggccc tggctcagag caacgagcag ggccaggtct gctaccagga gctggtggac   420

1

ctgatcagca gcaagcgctc cagcagtttc aagcgggcca ttgctaacgg acagcgggca   480

ctgcccccggg acgggccgct ggatgagcca ggcctaggtg tctacaagcg gtttgtgcgt   540

tacgtggcct acgagatcct gccttgtgag gtggaccgcc gctggtactt ctaccgtcac   600

cgcagctgcc caccccccgt gttcatggcc tcggtcactc ttgcccagat catcgtgttc   660

ctgtgttacg gggcccgcct caacaagtgg gtgctgcaga cctaccaccc cgagtacatg   720

aagagcccc ttgtgtacca ccccgggcac cgtgcccgcg cctggcgctt cctcacctac   780

atgttcatgc acgttgggct ggagcagctg gggttcaacg ccctcctgca gctgatgatc   840

ggggtgcccc tggagatggt gcacggcctg ctccgcatca gcctgctcta cctggcaggc   900

gtgctggcag gctccctaac cgtctccatc accgacatgc ggggccccggt ggtgggaggc   960

tccggcgggg tctacgccct gtgctcggca cacctggcca acgttgtcat gaactgggct   1020

gggatgagat gtccctacaa gttgctgagg atggtgctgg ccttggtgtg catgagctcc   1080

gaggtgggcc gggccgtgtg gctgcgcttc tccccgccgc tgcccgcctc gggcccacag   1140

cccagcttca tggcgcacct ggcaggcgcg gtggtggggg tgagcatggg cctgaccatc   1200

ctgcggagct acgaggagcg cctgcgggac cagtgcggct ggtgggtggt gctgctggcc   1260

tacggcacct tcctgctctt cgccgtcttc tggaacgtct cgcctacga cctgctgggc   1320

gcccacatcc ccccaccgcc ctgaccggct acctgaggct gcacaggcca gggctcgggc   1380

atgtggtggc cgcccaccag gggccttcac gtctgccctt tgtgaacgga cgtctcaggg   1440

ctgctgtgcc ccttgggtgt gggtggcctc aaaggaggcc ctgtcccagc cacccacccc   1500

ccactcccag gacttgcggt ctgagccttt ttggataatt aataaatatt ttacacagc   1559

<210> 2
<211> 438
<212> PRT
<213> Homo sapiens

<400> 2

Met Gly Arg Val Glu Asp Gly Gly Thr Thr Glu Glu Leu Glu Asp Trp
1   5   10   15

Asp Pro Gly Thr Ser Ala Leu Pro Ala Pro Gly Ile Lys Gln Gly Pro
            20   25   30

Arg Glu Gln Thr Gly Thr Gly Pro Leu Ser Gln Lys Cys Trp Glu Pro
        35   40   45

Glu Pro Asp Ala Pro Ser Gln Pro Gly Pro Ala Leu Trp Ser Arg Gly
    50   55   60

                                2

```
Arg Ala Arg Thr Gln Ala Leu Ala Gly Gly Ser Ser Leu Gln Gln Leu
65              70          75                  80


Asp Pro Glu Asn Thr Gly Phe Ile Gly Ala Asp Thr Phe Thr Gly Leu
            85              90              95


Val His Ser His Glu Leu Pro Leu Asp Pro Ala Lys Leu Asp Met Leu
        100             105             110


Val Ala Leu Ala Gln Ser Asn Glu Gln Gly Gln Val Cys Tyr Gln Glu
        115             120             125


Leu Val Asp Leu Ile Ser Ser Lys Arg Ser Ser Ser Phe Lys Arg Ala
    130             135             140


Ile Ala Asn Gly Gln Arg Ala Leu Pro Arg Asp Gly Pro Leu Asp Glu
145             150             155             160


Pro Gly Leu Gly Val Tyr Lys Arg Phe Val Arg Tyr Val Ala Tyr Glu
            165             170             175


Ile Leu Pro Cys Glu Val Asp Arg Arg Trp Tyr Phe Tyr Arg His Arg
            180             185             190


Ser Cys Pro Pro Pro Val Phe Met Ala Ser Val Thr Leu Ala Gln Ile
        195             200             205


Ile Val Phe Leu Cys Tyr Gly Ala Arg Leu Asn Lys Trp Val Leu Gln
    210             215             220


Thr Tyr His Pro Glu Tyr Met Lys Ser Pro Leu Val Tyr His Pro Gly
225             230             235             240


His Arg Ala Arg Ala Trp Arg Phe Leu Thr Tyr Met Phe Met His Val
            245             250             255


Gly Leu Glu Gln Leu Gly Phe Asn Ala Leu Leu Gln Leu Met Ile Gly
            260             265             270


Val Pro Leu Glu Met Val His Gly Leu Leu Arg Ile Ser Leu Leu Tyr
    275             280             285


Leu Ala Gly Val Leu Ala Gly Ser Leu Thr Val Ser Ile Thr Asp Met
    290             295             300


Arg Ala Pro Val Val Gly Gly Ser Gly Gly Val Tyr Ala Leu Cys Ser
305             310             315             320
```

                                    3

```
Ala His Leu Ala Asn Val Val Met Asn Trp Ala Gly Met Arg Cys Pro
                325                 330                 335


Tyr Lys Leu Leu Arg Met Val Leu Ala Leu Val Cys Met Ser Ser Glu
                340                 345                 350


Val Gly Arg Ala Val Trp Leu Arg Phe Ser Pro Pro Leu Pro Ala Ser
                355                 360                 365


Gly Pro Gln Pro Ser Phe Met Ala His Leu Ala Gly Ala Val Val Gly
        370             · 375                 380


Val Ser Met Gly Leu Thr Ile Leu Arg Ser Tyr Glu Glu Arg Leu Arg
385                 390         ·       395                 400


Asp Gln Cys Gly Trp Trp Val Val Leu Leu Ala Tyr Gly Thr Phe Leu
                405                 410                 415


Leu Phe Ala Val Phe Trp Asn Val Phe Ala Tyr Asp Leu Leu Gly Ala
                420                 425     ·           430


His Ile Pro Pro Pro Pro
            435
```

<210> 3
<211> 1224
<212> DNA
<213> Homo sapiens

<400> 3

```
gggccattta ggaggtttag atcattttga tcatcttcag ctgtcttctc ttcacataca      60

ggaaaggcct tggaaagcag tcgttgcgcc agacagccca gggaagagcg gcagcctgag     120

gacctagggc cacctgctgt tccctgggat tcatgtcctt ctggggagga gggaggaccc     180

aggacaatgg ctgctgttca tgatctggag atggagagca tgaatctgaa tatggggaga     240

gagatgaaag aagagctgga ggaagaggag aaaatgagag aggatggggg aggtaaagat     300

cgggccaaga gtaaaaaggt ccacaggatt gtctcaaaat ggatgctgcc cgaaaagtcc     360

cgaggaacat acttggagag agctaactgc ttcccgcctc ccgtgttcat catctccatc     420

agcctggccg agctggcagt gtttatttac tatgctgtgt ggaagcctca gaaacagtgg     480

atcacgttgg acacaggcat cttggagagt ccctttatct acagtcctga gaagagggag     540

gaagcctgga ggtttatctc atacatgctg gtacatgctg gagttcagca catcttgggg     600

aatctttgta tgcagcttgt tttgggtatt cccttggaaa tggtccacaa aggcctccgt     660
```

                                        4

```
gtggggctgg tgtacctggc aggagtgatt gcagggtccc ttgccagctc catctttgac      720

ccactcagat atcttgtggg agcttcagga ggagtctatg ctctgatggg aggctatttt      780

atgaatgttc tggtgaattt tcaagaaatg attcctgcct ttggaatttt cagactgctg      840

atcatcatcc tgataattgt gttggacatg ggatttgctc tctatagaag gttctttgtt      900

cctgaagatg ggtctccggt gtcttttgca gctcacattg caggtggatt tgctggaatg      960

tccattggct acacggtgtt tagctgcttt gataaagcac tgctgaaaga tccaaggttt     1020

tggatagcaa ttgctgcata tttagcttgt gtcttatttg ctgtgttttt caacattttc     1080

ctatctccag caaactgacc tgcccctatt gtaagtcaat taataaaaag agccatctgg     1140

aggaaataaa aaaaaaagga agactctatg aagaaacaga gaagtctcag aaaaggctaa     1200

caattttaga tagagaacaa aggg                                            1224
```

<210> 4
<211> 303
<212> PRT
<213> Homo sapiens

<400> 4

```
Met Ala Ala Val His Asp Leu Glu Met Glu Ser Met Asn Leu Asn Met
1               5               10              15


Gly Arg Glu Met Lys Glu Glu Leu Glu Glu Glu Glu Lys Met Arg Glu
            20              25              30


Asp Gly Gly Gly Lys Asp Arg Ala Lys Ser Lys Lys Val His Arg Ile
        35              40              45


Val Ser Lys Trp Met Leu Pro Glu Lys Ser Arg Gly Thr Tyr Leu Glu
    50              55              60


Arg Ala Asn Cys Phe Pro Pro Val Phe Ile Ile Ser Ile Ser Leu
65              70              75              80


Ala Glu Leu Ala Val Phe Ile Tyr Tyr Ala Val Trp Lys Pro Gln Lys
                85              90              95


Gln Trp Ile Thr Leu Asp Thr Gly Ile Leu Glu Ser Pro Phe Ile Tyr
            100             105             110


Ser Pro Glu Lys Arg Glu Glu Ala Trp Arg Phe Ile Ser Tyr Met Leu
        115             120             125


Val His Ala Gly Val Gln His Ile Leu Gly Asn Leu Cys Met Gln Leu
    130             135             140
```

5

```
Val Leu Gly Ile Pro Leu Glu Met Val His Lys Gly Leu Arg Val Gly
145             150             155             160

Leu Val Tyr Leu Ala Gly Val Ile Ala Gly Ser Leu Ala Ser Ser Ile
            165             170             175

Phe Asp Pro Leu Arg Tyr Leu Val Gly Ala Ser Gly Gly Val Tyr Ala
        180             185             190

Leu Met Gly Gly Tyr Phe Met Asn Val Leu Val Asn Phe Gln Glu Met
        195             200             205

Ile Pro Ala Phe Gly Ile Phe Arg Leu Leu Ile Ile Ile Leu Ile Ile
    210             215             220

Val Leu Asp Met Gly Phe Ala Leu Tyr Arg Arg Phe Phe Val Pro Glu
225             230             235             240

Asp Gly Ser Pro Val Ser Phe Ala Ala His Ile Ala Gly Gly Phe Ala
            245             250             255

Gly Met Ser Ile Gly Tyr Thr Val Phe Ser Cys Phe Asp Lys Ala Leu
            260             265             270

Leu Lys Asp Pro Arg Phe Trp Ile Ala Ile Ala Ala Tyr Leu Ala Cys
        275             280             285

Val Leu Phe Ala Val Phe Phe Asn Ile Phe Leu Ser Pro Ala Asn
    290             295             300
```

```
<210>   5
<211>   1376
<212>   DNA
<213>   Homo sapiens

<400>   5
ctgaactgat gaggttttca tgggaaatca gtgcctgtct gttcatctga acacccacta        60

gttattcatc caacaaatct tgattgagtg ctgataatgc caggctctgt gctaagtacc       120

ggggataagt tgtgataccg acccgcgagg cgccgcggtc caaggaggag gcaaaagcag       180

acagacatca gtgtgggctg gaggcccaga ggtctggaca gaacagaggg ttccgtgaga       240

acaggccatg gctgagtttg accctgggaa cacaggctac attagcacag caagttccg        300

gagtcttctg gagagccaca gctccaagct ggacccgcac aaaagggagg tcctcctggc       360

tcttgccgac agccacgcgg atgggcagat cggctaccag gattttgtca gcctagtgag       420
```

```
caacaagcgt tccaacagct tccgccaagc catcctgcag ggcaaccgca ggctaagcag     480

caaggccctg ctggaggaga aggggctgag cctctcgcag cgacttatcc gccatgtggc     540

ctatgagacc ctgccccggg aaattgaccg caagtggtac tatgacagct acacctgctg     600

cccccacccc tggttcatga tcacagtcac gctgctggag gttgcctttt tcctctacaa     660

tggggtgtca ctaggtcaat ttgtactgca ggtaactcat ccacgttact tgaagaactc     720

cctggtttac cacccacagc tgcgagcaca ggtttggcgc tacctgacat acatcttcat     780

gcatgcaggg atagaacacc tgggactcaa tgtggtgctg cagctgctgg tgggggtgcc     840

cctggagatg gtgcatggag ccacccgaat tgggcttgtc tacgtggccg gtgttgtggc     900

agggtccttg gcagtgtctg tggctgacat gaccgctcca gtcgtgggct cttctggagg     960

ggtgtatgct ctcgtctctg cccatctggc caacattgtc atgaactggt caggcatgaa    1020

gtgccagttc aagctgctgc ggatggctgt ggcccttatc tgtatgagca tggagtttgg    1080

gcgggccgtg tggctccgct ccacccgtc ggcctatccc ccgtgccctc acccaagctt    1140

tgtggcgcac ttgggtggcg tggccgtggg catcaccctg ggcgtggtgg tcctgaggaa    1200

ctacgagcag aggctccagg accagtcact gtggtggatt tttgtggcca tgtacaccgt    1260

cttcgtgctg ttcgctgtct tctggaacat ctttgcctac accctgctgg acttaaagct    1320

gccgcctccc ccctgagggc tggaggccca aggtcgggga ggggagggaa aagcag       1376
```

```
<210>   6
<211>   362
<212>   PRT
<213>   Homo sapiens

<400>   6
```

```
Met Ala Glu Phe Asp Pro Gly Asn Thr Gly Tyr Ile Ser Thr Gly Lys
1               5                   10                  15


Phe Arg Ser Leu Leu Glu Ser His Ser Ser Lys Leu Asp Pro His Lys
            20                  25                  30


Arg Glu Val Leu Leu Ala Leu Ala Asp Ser His Ala Asp Gly Gln Ile
        35                  40                  45


Gly Tyr Gln Asp Phe Val Ser Leu Val Ser Asn Lys Arg Ser Asn Ser
    50                  55                  60


Phe Arg Gln Ala Ile Leu Gln Gly Asn Arg Arg Leu Ser Ser Lys Ala
65                  70                  75                  80


Leu Leu Glu Glu Lys Gly Leu Ser Leu Ser Gln Arg Leu Ile Arg His
                85                  90                  95
```

Val Ala Tyr Glu Thr Leu Pro Arg Glu Ile Asp Arg Lys Trp Tyr Tyr
            100   105   110

Asp Ser Tyr Thr Cys Cys Pro Pro Pro Trp Phe Met Ile Thr Val Thr
            115   120   125

Leu Leu Glu Val Ala Phe Phe Leu Tyr Asn Gly Val Ser Leu Gly Gln
    130   135   140

Phe Val Leu Gln Val Thr His Pro Arg Tyr Leu Lys Asn Ser Leu Val
145   150   155   160

Tyr His Pro Gln Leu Arg Ala Gln Val Trp Arg Tyr Leu Thr Tyr Ile
                165   170   175

Phe Met His Ala Gly Ile Glu His Leu Gly Leu Asn Val Val Leu Gln
            180   185   190

Leu Leu Val Gly Val Pro Leu Glu Met Val His Gly Ala Thr Arg Ile
    195   200   205

Gly Leu Val Tyr Val Ala Gly Val Val Ala Gly Ser Leu Ala Val Ser
    210   215   220

Val Ala Asp Met Thr Ala Pro Val Val Gly Ser Ser Gly Gly Val Tyr
225   230   235   240

Ala Leu Val Ser Ala His Leu Ala Asn Ile Val Met Asn Trp Ser Gly
            245   250   255

Met Lys Cys Gln Phe Lys Leu Leu Arg Met Ala Val Ala Leu Ile Cys
            260   265   270

Met Ser Met Glu Phe Gly Arg Ala Val Trp Leu Arg Phe His Pro Ser
    275   280   285

Ala Tyr Pro Pro Cys Pro His Pro Ser Phe Val Ala His Leu Gly Gly
    290   295   300

Val Ala Val Gly Ile Thr Leu Gly Val Val Val Leu Arg Asn Tyr Glu
305   310   315   320

Gln Arg Leu Gln Asp Gln Ser Leu Trp Trp Ile Phe Val Ala Met Tyr
                325   330   335

8

```
Thr Val Phe Val Leu Phe Ala Val Phe Trp Asn Ile Phe Ala Tyr Thr
            340                 345                 350


Leu Leu Asp Leu Lys Leu Pro Pro Pro Pro
            355                 360
```

<210> 7
<211> 1140
<212> DNA
<213> Homo sapiens

<400> 7

```
atggcgtggc gaggctgggc gcagagaggc tggggctgcg gccaggcgtg gggtgcgtcg      60

gtgggcggcc gcagctgcga ggagctcact gcggtcctaa ccccgccgca gctcctcgga     120

cgcaggttta acttctttat tcaacaaaaa tgcggattca gaaaagcacc caggaaggtt     180

gaacctcgaa gatcagaccc agggacaagt ggtgaagcat acaagagaag tgctttgatt     240

cctcctgtgg aagaaacagt cttttatcct tctccctatc ctataaggag tctcataaaa     300

cctttatttt ttactgttgg gtttacaggc tgtgcatttg atcagctgc tatttggcaa      360

tatgaatcac tgaaatccag ggtccagagt tattttgatg gtataaaagc tgattggttg     420

gatagcataa gaccacaaaa agaaggagac ttcagaaagg agattaacaa gtggtggaat     480

aacctaagtg atggccagcg gactgtgaca ggtattatag ctgcaaatgt ccttgtattc     540

tgtttatgga gagtaccttc tctgcagcgg acaatgatca gatatttcac atcgaatcca     600

gcctcaaagg tcctttgttc tccaatgttg ctgtcaacat tcagtcactt ctccttattt     660

cacatggcag caaatatgta tgttttgtgg agcttctctt ccagcatagt gaacattctg     720

ggtcaagagc agttcatggc agtgtaccta tctgcaggtg ttatttccaa ttttgtcagt     780

tacctgggta aagttgccac aggaagatat ggaccatcac ttggtgcatc tggtgccatc     840

atgacagtcc tcgcagctgt ctgcactaag atcccagaag ggaggcttgc cattattttc     900

cttccgatgt tcacgttcac agcagggaat gccctgaaag ccattatcgc catggataca     960

gcaggaatga tcctgggatg gaaatttttt gatcatgcgg cacatcttgg gggagctctt    1020

tttggaatat ggtatgttac ttacggtcat gaactgattt ggaagaacag ggagccgcta    1080

gtgaaaatct ggcatgaaat aaggactaat ggccccaaaa aggaggtgg ctctaagtaa     1140
```

<210> 8
<211> 379
<212> PRT
<213> Homo sapiens

<400> 8

```
Met Ala Trp Arg Gly Trp Ala Gln Arg Gly Trp Gly Cys Gly Gln Ala
1               5                   10                  15
```

9

Trp Gly Ala Ser Val Gly Gly Arg Ser Cys Glu Glu Leu Thr Ala Val
            20   25   30

Leu Thr Pro Pro Gln Leu Leu Gly Arg Arg Phe Asn Phe Phe Ile Gln
            35   40   45

Gln Lys Cys Gly Phe Arg Lys Ala Pro Arg Lys Val Glu Pro Arg Arg
    50   55   60

Ser Asp Pro Gly Thr Ser Gly Glu Ala Tyr Lys Arg Ser Ala Leu Ile
65   70   75   80

Pro Pro Val Glu Glu Thr Val Phe Tyr Pro Ser Pro Tyr Pro Ile Arg
                85   90   95

Ser Leu Ile Lys Pro Leu Phe Phe Thr Val Gly Phe Thr Gly Cys Ala
            100   105   110

Phe Gly Ser Ala Ala Ile Trp Gln Tyr Glu Ser Leu Lys Ser Arg Val
            115   120   125

Gln Ser Tyr Phe Asp Gly Ile Lys Ala Asp Trp Leu Asp Ser Ile Arg
    130   135   140

Pro Gln Lys Glu Gly Asp Phe Arg Lys Glu Ile Asn Lys Trp Trp Asn
145   150   155   160

Asn Leu Ser Asp Gly Gln Arg Thr Val Thr Gly Ile Ile Ala Ala Asn
            165   170   175

Val Leu Val Phe Cys Leu Trp Arg Val Pro Ser Leu Gln Arg Thr Met
            180   185   190

Ile Arg Tyr Phe Thr Ser Asn Pro Ala Ser Lys Val Leu Cys Ser Pro
    195   200   205

Met Leu Leu Ser Thr Phe Ser His Phe Ser Leu Phe His Met Ala Ala
    210   215   220

Asn Met Tyr Val Leu Trp Ser Phe Ser Ser Ser Ile Val Asn Ile Leu
225   230   235   240

Gly Gln Glu Gln Phe Met Ala Val Tyr Leu Ser Ala Gly Val Ile Ser
                245   250   255

10

Asn Phe Val Ser Tyr Leu Gly Lys Val Ala Thr Gly Arg Tyr Gly Pro
            260   265   270

Ser Leu Gly Ala Ser Gly Ala Ile Met Thr Val Leu Ala Ala Val Cys
            275   280   285

Thr Lys Ile Pro Glu Gly Arg Leu Ala Ile Ile Phe Leu Pro Met Phe
        290   295   300

Thr Phe Thr Ala Gly Asn Ala Leu Lys Ala Ile Ile Ala Met Asp Thr
305   310   315   320

Ala Gly Met Ile Leu Gly Trp Lys Phe Phe Asp His Ala Ala His Leu
                325   330   335

Gly Gly Ala Leu Phe Gly Ile Trp Tyr Val Thr Tyr Gly His Glu Leu
            340   345   350

Ile Trp Lys Asn Arg Glu Pro Leu Val Lys Ile Trp His Glu Ile Arg
            355   360   365

Thr Asn Gly Pro Lys Lys Gly Gly Gly Ser Lys
        370   375

<210>   9
<211>   2964
<212>   DNA
<213>   Homo sapiens

<400>   9
gactcggcgc gggcgccctc ccggccagcg gcggcagccc ctcctccccg gcgccctcag   60

gacccccccag agaccccccgg cggcggcagc ctgccttgct ctgccaggaa ccatgagtga   120

ggcccgcagg gacagcacga gcagcctgca cgcaagaag ccaccctggc taaagctgga   180

cattccctct gcggtgcccc tgacggcaga agagcccagc ttcctgcagc ccctgaggcg   240

acaggctttc ctgaggagtg tgagtatgcc agccgagaca gcccacatct cttcacccca   300

ccatgagctc cggcggccgg tgctgcaacg ccagacgtcc atcacacaga ccatccgcag   360

ggggaccgcc gactggtttg gagtgagcaa ggacagtgac agcacccaga aatggcagcg   420

caagagcatc cgtcactgca gccagcgcta cgggaagctg aagccccagg tcctccggga   480

gctggacctg cccagccagg acaacgtgtc gctgaccagc accgagacgc cacccccact   540

ctacgtgggg ccatgccagc tgggcatgca gaagatcata gacccccctgg cccgtggccg   600

tgccttccgt gtggcagatg acactgcgga aggcctgagt gccccacaca ctcccgtcac   660

gccgggtgct gcctccctct gctccttctc cagctcccgc tcaggtttcc accggctccc   720

11

```
gcggcggcgc aagcgagagt cggtggccaa gatgagcttc cgggcggccg cagcgctgat      780

gaaaggccgc tccgttaggg atggcacctt tcgccgggca cagcgtcgaa gcttcactcc      840

agctagcttt ctggaggagg acacaactga tttccccgat gagctggaca catccttctt      900

tgcccgggaa ggtatcctcc atgaagagct gtccacatac ccggatgaag ttttcgagtc      960

cccatcggag gcagcgctaa aggactggga gaaggcaccg gagcaggcgg acctcaccgg     1020

cggggccctg gaccgcagcg agcttgagcg cagccacctg atgctgccct ggagcgagg      1080

ctggcggaag cagaaggagg gcgccgcagc cccgcagccc aaggtgcggc tccgacagga     1140

ggtggtgagc accgcggggc cgcgacgggg ccagcgtatc gcggtgccgg tgcgcaagct     1200

cttcgcccgg gagaagcggc cgtatgggct gggcatggtg ggacggctca ccaaccgcac     1260

ctaccgcaag cgcatcgaca gcttcgtcaa cgccagatc gaggacatgg acgaccacag     1320

gcccttcttc acctactggc ttaccttcgt gcactcgctc gtcgccatcc tagccgtgtg     1380

catctatggc atcgcgcccg tgggcttctc gcagcatgag acggtggact cggtgctgcg     1440

gaaccgcggg gtctacgaga acgtcaagta cgtgcagcag gagaacttct ggatcgggcc     1500

cagctcggag gccctcatcc acctgggcgc caagttttcg ccctgcatgc gccaggaccc     1560

gcaggtgcac agcttcattc gctcggcgcg cgagcgcgag aagcactccg cctgctgcgt     1620

gcgcaacgac aggtcgggct gcgtgcagac ctcggaggag gagtgctcgt ccacgctggc     1680

agtgtgggtg aagtggccca tccatcccag cgccccagag cttgcgggcc acaagagaca     1740

gtttggctct gtctgccacc aggatcccag ggtgtgtgat gagccctcct ccgaagaccc     1800

tcatgagtgg ccagaagaca tcaccaagtg gccgatctgc accaaaaaca gcgctgggaa     1860

ccacaccaac catccccaca tggactgtgt catcacagga cggccctgct gcattggcac     1920

caagggcagg tgtgagatca cctcccggga gtactgtgac ttcatgaggg gctacttcca     1980

tgaggaggcc acgctctgct ctcaggtgca ctgcatggat gatgtgtgtg ggctcctgcc     2040

ttttctcaac cccgaggtgc ctgaccagtt ctaccgcctg tggctatccc tcttcctgca     2100

cgccgggatc ttgcactgcc tggtgtccat ctgcttccag atgactgtcc tgcgggacct     2160

ggagaagctg gcaggctggc accgcatagc catcatctac ctgctgagtg gtgtcaccgg     2220

caacctggcc agtgccatct tcctgccata ccgagcagag gtgggtcctg ctggctccca     2280

gttcggcatc ctggcctgcc tcttcgtgga gctcttccag agctggcaga tcctggcgcg     2340

gccctggcgt gccttcttca gctgctggc tgtggtgctc ttcctcttca cctttgggct     2400

gctgccgtgg attgacaact ttgcccacat ctcggggttc atcagtggcc tcttcctctc     2460

cttcgccttc ttgccctaca tcagctttgg caagttcgac ctgtaccgga aacgctgcca     2520

gatcatcatc tttcaggtgg tcttcctggg cctcctggct ggcctggtgg tcctcttcta     2580
```

                                       12

cgtctatcct gtccgctgtg agtggtgtga gttcctcacc tgcatcccct tcactgacaa   2640

gttctgtgag aagtacgaac tggacgctca gctccactga gctggctgcg ggctccagcg   2700

gccgtgtgct ccagcaggcc agagccagac acgacctccc tgagcctcac aggcttacag   2760

gagtcacctg ctccatgtgg ggactggcct gtttctgaa cacagacctc tttcttgtgc   2820

cttgttcact tctgttgaac ccctcgtact gccgggcatt tattatacta cttcctgtca   2880

taaccttcta acttgtttct tgacgaccac ctcatgtggc caataaatga actgggagcg   2940

ttttaaaaaa aaaaaaaaaa aaaa   2964

<210> 10
<211> 855
<212> PRT
<213> Homo sapiens

<400> 10

Met Ser Glu Ala Arg Arg Asp Ser Thr Ser Ser Leu Gln Arg Lys Lys
1   5   10   15

Pro Pro Trp Leu Lys Leu Asp Ile Pro Ser Ala Val Pro Leu Thr Ala
            20   25   30

Glu Glu Pro Ser Phe Leu Gln Pro Leu Arg Arg Gln Ala Phe Leu Arg
        35   40   45

Ser Val Ser Met Pro Ala Glu Thr Ala His Ile Ser Ser Pro His His
    50   55   60

Glu Leu Arg Arg Pro Val Leu Gln Arg Gln Thr Ser Ile Thr Gln Thr
65   70   75   80

Ile Arg Arg Gly Thr Ala Asp Trp Phe Gly Val Ser Lys Asp Ser Asp
                85   90   95

Ser Thr Gln Lys Trp Gln Arg Lys Ser Ile Arg His Cys Ser Gln Arg
            100   105   110

Tyr Gly Lys Leu Lys Pro Gln Val Leu Arg Glu Leu Asp Leu Pro Ser
        115   120   125

Gln Asp Asn Val Ser Leu Thr Ser Thr Glu Thr Pro Pro Pro Leu Tyr
    130   135   140

Val Gly Pro Cys Gln Leu Gly Met Gln Lys Ile Ile Asp Pro Leu Ala
145   150   155   160

13

Arg Gly Arg Ala Phe Arg Val Ala Asp Asp Thr Ala Glu Gly Leu Ser
                165   170   175

Ala Pro His Thr Pro Val Thr Pro Gly Ala Ala Ser Leu Cys Ser Phe
            180   185   190

Ser Ser Ser Arg Ser Gly Phe His Arg Leu Pro Arg Arg Arg Lys Arg
        195   200   205

Glu Ser Val Ala Lys Met Ser Phe Arg Ala Ala Ala Ala Leu Met Lys
    210   215   220

Gly Arg Ser Val Arg Asp Gly Thr Phe Arg Arg Ala Gln Arg Arg Ser
225   230   235   240

Phe Thr Pro Ala Ser Phe Leu Glu Glu Asp Thr Thr Asp Phe Pro Asp
            245   250   255

Glu Leu Asp Thr Ser Phe Phe Ala Arg Glu Gly Ile Leu His Glu Glu
            260   265   270

Leu Ser Thr Tyr Pro Asp Glu Val Phe Glu Ser Pro Ser Glu Ala Ala
        275   280   285

Leu Lys Asp Trp Glu Lys Ala Pro Glu Gln Ala Asp Leu Thr Gly Gly
    290   295   300

Ala Leu Asp Arg Ser Glu Leu Glu Arg Ser His Leu Met Leu Pro Leu
305   310   315   320

Glu Arg Gly Trp Arg Lys Gln Lys Glu Gly Ala Ala Ala Pro Gln Pro
            325   330   335

Lys Val Arg Leu Arg Gln Glu Val Val Ser Thr Ala Gly Pro Arg Arg
        340   345   350

Gly Gln Arg Ile Ala Val Pro Val Arg Lys Leu Phe Ala Arg Glu Lys
        355   360   365

Arg Pro Tyr Gly Leu Gly Met Val Gly Arg Leu Thr Asn Arg Thr Tyr
    370   375   380

Arg Lys Arg Ile Asp Ser Phe Val Lys Arg Gln Ile Glu Asp Met Asp
385   390   395   400

Asp His Arg Pro Phe Phe Thr Tyr Trp Leu Thr Phe Val His Ser Leu
            405   410   415

14

```
Val Ala Ile Leu Ala Val Cys Ile Tyr Gly Ile Ala Pro Val Gly Phe
            420                 425             430

Ser Gln His Glu Thr Val Asp Ser Val Leu Arg Asn Arg Gly Val Tyr
        435                 440             445

Glu Asn Val Lys Tyr Val Gln Gln Glu Asn Phe Trp Ile Gly Pro Ser
    450                 455             460

Ser Glu Ala Leu Ile His Leu Gly Ala Lys Phe Ser Pro Cys Met Arg
465             470             475                         480

Gln Asp Pro Gln Val His Ser Phe Ile Arg Ser Ala Arg Glu Arg Glu
            485                 490             495

Lys His Ser Ala Cys Cys Val Arg Asn Asp Arg Ser Gly Cys Val Gln
            500             505             510

Thr Ser Glu Glu Glu Cys Ser Ser Thr Leu Ala Val Trp Val Lys Trp
        515             520             525

Pro Ile His Pro Ser Ala Pro Glu Leu Ala Gly His Lys Arg Gln Phe
    530             535             540

Gly Ser Val Cys His Gln Asp Pro Arg Val Cys Asp Glu Pro Ser Ser
545             550             555                         560

Glu Asp Pro His Glu Trp Pro Glu Asp Ile Thr Lys Trp Pro Ile Cys
            565             570             575

Thr Lys Asn Ser Ala Gly Asn His Thr Asn His Pro His Met Asp Cys
            580             585             590

Val Ile Thr Gly Arg Pro Cys Cys Ile Gly Thr Lys Gly Arg Cys Glu
    595             600             605

Ile Thr Ser Arg Glu Tyr Cys Asp Phe Met Arg Gly Tyr Phe His Glu
    610             615             620

Glu Ala Thr Leu Cys Ser Gln Val His Cys Met Asp Asp Val Cys Gly
625             630             635                         640

Leu Leu Pro Phe Leu Asn Pro Glu Val Pro Asp Gln Phe Tyr Arg Leu
            645             650             655
```

15

Trp Leu Ser Leu Phe Leu His Ala Gly Ile Leu His Cys Leu Val Ser
            660   665   670

Ile Cys Phe Gln Met Thr Val Leu Arg Asp Leu Glu Lys Leu Ala Gly
            675   680   685

Trp His Arg Ile Ala Ile Ile Tyr Leu Leu Ser Gly Val Thr Gly Asn
    690   695   700

Leu Ala Ser Ala Ile Phe Leu Pro Tyr Arg Ala Glu Val Gly Pro Ala
705   710   715   720

Gly Ser Gln Phe Gly Ile Leu Ala Cys Leu Phe Val Glu Leu Phe Gln
            725   730   735

Ser Trp Gln Ile Leu Ala Arg Pro Trp Arg Ala Phe Phe Lys Leu Leu
            740   745   750

Ala Val Val Leu Phe Leu Phe Thr Phe Gly Leu Leu Pro Trp Ile Asp
            755   760   765

Asn Phe Ala His Ile Ser Gly Phe Ile Ser Gly Leu Phe Leu Ser Phe
    770   775   780

Ala Phe Leu Pro Tyr Ile Ser Phe Gly Lys Phe Asp Leu Tyr Arg Lys
785   790   795   800

Arg Cys Gln Ile Ile Ile Phe Gln Val Val Phe Leu Gly Leu Leu Ala
            805   810   815

Gly Leu Val Val Leu Phe Tyr Val Tyr Pro Val Arg Cys Glu Trp Cys
            820   825   830

Glu Phe Leu Thr Cys Ile Pro Phe Thr Asp Lys Phe Cys Glu Lys Tyr
            835   840   845

Glu Leu Asp Ala Gln Leu His
    850   855

<210>   11
<211>   2643
<212>   DNA
<213>   Homo sapiens

<400>   11
gctactccca cctgccacgc cgcaagagaa tgtctgtggc ccacatgagc ttgcaagctg   60

ccgctgccct cctcaagggg cgctcggtgc tggatgccac cggacagcgg tgccgggtgg   120

16

```
tcaagcgcag ctttgccttc ccgagcttcc tggaggagga tgtggtcgat ggggcagaca   180

cgtttgactc ctcctttttt agtaaggaag aaatgagctc catgcctgat gatgtctttg   240

agtcccccccc actctctgcc agctacttcc gagggatccc acactcagcc tcccctgtct   300

cccccgatgg ggtgcaaatc cctctgaagg agtatggccg agccccagtc cccgggcccc   360

ggcgcggcaa gcgcatcgcc tccaaggtga agcactttgc ctttgatcgg aagaagcggc   420

actacggcct cggcgtggtg ggcaactggc tgaaccgcag ctaccgccgc agcatcagca   480

gcactgtgca gcggcagctg gagagcttcg acagccaccg gccctacttc acctactggc   540

tgaccttcgt ccatgtcatc atcacgctgc tggtgatttg cacgtatggc atcgcacccg   600

tgggctttgc ccagcacgtc accacccagc tggtgctgcg gaacaaaggt gtgtacgaga   660

gcgtgaagta catccagcag gagaacttct gggttggccc cagctcgatt gacctgatcc   720

acctggggggc caagttctca ccctgcatcc ggaaggacgg gcagatcgag cagctggtgc   780

tgcgcgagcg agacctggag cgggactcag gctgctgtgt ccagaatgac cactccggat   840

gcatccagac ccagcggaag gactgctcgg gactttggc cacttttgtc aagtggcagg   900

atgacactgg gccccccatg gacaagtctg atctgggcca gaagcggact tcggggggctg   960

tctgccacca ggaccccagg acctgcgagg agccagcctc cagcggtgcc cacatctggc  1020

ccgatgacat cactaagtgg ccgatctgca cagagcaggc caggagcaac cacacaggct  1080

tcctgcacat ggactgcgag atcaagggcc gccctgctg catcggcacc aagggcagct  1140

gtgagatcac cacccgggaa tactgtgagt tcatgcacgg ctatttccat gaggaagcaa  1200

cactctgctc ccaggtgcac tgcttggaca aggtgtgtgg gctgctgccc ttcctcaacc  1260

ctgaggtccc agatcagttc tacaggctct ggctgtctct cttcctacat gctggcgtgg  1320

tgcactgcct cgtgtctgtg gtctttcaaa tgaccatcct gagggacctg gagaagctgg  1380

ccggctggca ccgtatcgcc atcatcttca tcctcagtgg catcacaggc aacctcgcca  1440

gtgccatctt tctcccatac cgggcagagg tgggcccggc cggctcacag ttcggcctcc  1500

tcgcctgcct cttcgtggag ctcttccaga gctggccgct gctggagagg ccctggaagg  1560

ccttcctcaa cctctcggcc atcgtgctct cctgttcat ctgtggcctc ttgccctgga  1620

tcgacaacat cgcccacatc ttcggcttcc tcagtggcct gctgctggcc ttcgccttcc  1680

tgccctacat caccttcggc accagcgaca gtaccgcaa gcgggcactc atcctggtgt  1740

cactgctggc ctttgccggc ctcttcgccg ccctcgtgct gtggctgtac atctacccca  1800

ttaactggcc ctggatcgag cacctcacct gcttcccctt caccagccgc ttctgcgaga  1860

agtatgagct ggaccaggtg ctgcactgac cgctgggcca cacggctgcc cctcagccct  1920

gctggaacag ggtctgcctg cgagggctgc cctctgcaga gcgctctctg tgtgccagag  1980
```

17

```
agccagagac ccaagacagg gcccgggctc tggacctggg tgcccccctg ccaggcgagg    2040

ctgactccgc gtgagatggt tggttaaggc ggggtttttc tggggcgtga ggcctgtgag    2100

atcctgaccc aagctcaggc acacccaagg cacctgcctc tctgagtctt gggtctcagt    2160

tcctaatatc ccgctccttg ctgagaccat ctcctggggc agggtccttt tcttcccagg    2220

tcctcagcgc tgcctctgct ggtgccttct cccccactac tactggagcg tgcccttgct    2280

ggggacgtgg ctgtgccctc agttgccccc agggctgggt gcccaccatg ccccttcctc    2340

tttctcctcc tacctctgcc ctgtgagccc atccataagg ctctcagatg ggacattgta    2400

ggaaaggctt tggccatggt ctgggggcag agaacaaggg gggagacaca agtagacctc    2460

aggtagaacg acaccgggcg gagccacccc agggcctgct cccagggagt gctcgaggcg    2520

catcaggccc gttttttacc agtttatatc acggtcttca tttttaaaag taacgctaac    2580

tttgtacgga cgatgtctca tggattaaat aatattcttt atggcaaaaa aaaaaaaaaa    2640

aaa                                                                   2643
```

```
<210>  12
<211>  619
<212>  PRT
<213>  Homo sapiens

<400>  12

Met Ser Val Ala His Met Ser Leu Gln Ala Ala Ala Leu Leu Lys
1               5                   10                  15


Gly Arg Ser Val Leu Asp Ala Thr Gly Gln Arg Cys Arg Val Val Lys
            20                  25                  30


Arg Ser Phe Ala Phe Pro Ser Phe Leu Glu Glu Asp Val Val Asp Gly
        35                  40                  45


Ala Asp Thr Phe Asp Ser Ser Phe Phe Ser Lys Glu Glu Met Ser Ser
    50                  55                  60


Met Pro Asp Asp Val Phe Glu Ser Pro Pro Leu Ser Ala Ser Tyr Phe
65                  70                  75                  80


Arg Gly Ile Pro His Ser Ala Ser Pro Val Ser Pro Asp Gly Val Gln
                85                  90                  95


Ile Pro Leu Lys Glu Tyr Gly Arg Ala Pro Val Pro Gly Pro Arg Arg
                100                 105                 110


Gly Lys Arg Ile Ala Ser Lys Val Lys His Phe Ala Phe Asp Arg Lys
            115                 120                 125
```

18

Lys Arg His Tyr Gly Leu Gly Val Val Gly Asn Trp Leu Asn Arg Ser
    130   135   140

Tyr Arg Arg Ser Ile Ser Ser Thr Val Gln Arg Gln Leu Glu Ser Phe
145   150   155   160

Asp Ser His Arg Pro Tyr Phe Thr Tyr Trp Leu Thr Phe Val His Val
                165   170   175

Ile Ile Thr Leu Leu Val Ile Cys Thr Tyr Gly Ile Ala Pro Val Gly
            180   185   190

Phe Ala Gln His Val Thr Thr Gln Leu Val Leu Arg Asn Lys Gly Val
            195   200   205

Tyr Glu Ser Val Lys Tyr Ile Gln Gln Glu Asn Phe Trp Val Gly Pro
    210   215   220

Ser Ser Ile Asp Leu Ile His Leu Gly Ala Lys Phe Ser Pro Cys Ile
225   230   235   240

Arg Lys Asp Gly Gln Ile Glu Gln Leu Val Leu Arg Glu Arg Asp Leu
            245   250   255

Glu Arg Asp Ser Gly Cys Cys Val Gln Asn Asp His Ser Gly Cys Ile
            260   265   270

Gln Thr Gln Arg Lys Asp Cys Ser Glu Thr Leu Ala Thr Phe Val Lys
    275   280   285

Trp Gln Asp Asp Thr Gly Pro Pro Met Asp Lys Ser Asp Leu Gly Gln
    290   295   300

Lys Arg Thr Ser Gly Ala Val Cys His Gln Asp Pro Arg Thr Cys Glu
305   310   315   320

Glu Pro Ala Ser Ser Gly Ala His Ile Trp Pro Asp Asp Ile Thr Lys
                325   330   335

Trp Pro Ile Cys Thr Glu Gln Ala Arg Ser Asn His Thr Gly Phe Leu
            340   345   350

His Met Asp Cys Glu Ile Lys Gly Arg Pro Cys Cys Ile Gly Thr Lys
    355   360   365

```
Gly Ser Cys Glu Ile Thr Thr Arg Glu Tyr Cys Glu Phe Met His Gly
    370                 375             380


Tyr Phe His Glu Glu Ala Thr Leu Cys Ser Gln Val His Cys Leu Asp
385                 390             395                     400


Lys Val Cys Gly Leu Leu Pro Phe Leu Asn Pro Glu Val Pro Asp Gln
                405             410             415


Phe Tyr Arg Leu Trp Leu Ser Leu Phe Leu His Ala Gly Val Val His
            420             425             430


Cys Leu Val Ser Val Val Phe Gln Met Thr Ile Leu Arg Asp Leu Glu
        435             440             445


Lys Leu Ala Gly Trp His Arg Ile Ala Ile Ile Phe Ile Leu Ser Gly
    450             455             460


Ile Thr Gly Asn Leu Ala Ser Ala Ile Phe Leu Pro Tyr Arg Ala Glu
465             470             475                     480


Val Gly Pro Ala Gly Ser Gln Phe Gly Leu Leu Ala Cys Leu Phe Val
            485             490             495


Glu Leu Phe Gln Ser Trp Pro Leu Leu Glu Arg Pro Trp Lys Ala Phe
        500             505             510


Leu Asn Leu Ser Ala Ile Val Leu Phe Leu Phe Ile Cys Gly Leu Leu
        515             520             525


Pro Trp Ile Asp Asn Ile Ala His Ile Phe Gly Phe Leu Ser Gly Leu
    530             535             540


Leu Leu Ala Phe Ala Phe Leu Pro Tyr Ile Thr Phe Gly Thr Ser Asp
545             550             555                     560


Lys Tyr Arg Lys Arg Ala Leu Ile Leu Val Ser Leu Leu Ala Phe Ala
            565             570             575


Gly Leu Phe Ala Ala Leu Val Leu Trp Leu Tyr Ile Tyr Pro Ile Asn
            580             585             590


Trp Pro Trp Ile Glu His Leu Thr Cys Phe Pro Phe Thr Ser Arg Phe
    595             600             605


Cys Glu Lys Tyr Glu Leu Asp Gln Val Leu His
    610             615
```

20

<210> 13
<211> 1305
<212> DNA
<213> Homo sapiens

<400> 13
atgggcgggg gcctctggga ggcgtggcct ccggccggct cctctgctgt tgccaaggga   60

aactgccgcg aggaggcgga aggagcagag gaccggcagc cggcgtcgag gcggggcgcg   120

ggaacgacgg cggccatggc ggcctcgggg cccgggtgtc gcagctggtg cttgtgtccc   180

gaggtgccat ccgccacctt cttcactgcg ctgctctcgc tgctggtttc cgggcctcgc   240

ctgttcctgc tgcagcagcc cctggcgccc tcgggcctca cgctgaagtc cgaggccctt   300

cgcaactggc aagtttacag gctggtaacc tacatctttg tctacgagaa tcccatctcc   360

ctgctctgcg gcgctatcat catctggcgc tttgctggca atttcgagag aaccgtgggc   420

accgtccgcc actgcttctt caccgtgatc ttcgccatct tctccgctat catcttcctg   480

tcattcgagg ctgtgtcatc actgtcaaag ctgggggaag tggaggatgc cagaggtttc   540

accccagtgg cctttgccat gctgggagtc accaccgtcc gttctcggat gaggcgggcc   600

ctggtgtttg gcatggttgt gccctcagtc ctggttccgt ggctcctgct gggtgcctcg   660

tggctcattc cccagacctc tttcctcagt aatgtctgcg ggctgtccat cgggctggcc   720

tatgctcacc tactgctatt ccatcgacct ctcagagcga gtggcgctga agctcgatca   780

gaccttcccc ttcagcctga tgaggaggat atccgtgttc aagtacgtct cagggtcttc   840

agccgagagg agggcagccc agagccggaa actgaacccg gtgcctggct cctaccccac   900

acagagctgc caccctcacc tgtccccaag ccaccctgtg tcccagacgc agcacgccag   960

tggtcagaag ctggcctcct ggcctcctgc acccccgggc acatgcccac cttgcctccg   1020

taccagcctg cctccggcct gtgctatgtg cagaaccact ttggtccaaa ccccacctcc   1080

tccagtgtct acccagcttc tgcgggcacc tccctgggca tccagccccc cacgcctgtg   1140

aacagccctg cacggtgta ttctggggcc ttgggacacc aggggctgca ggctccaagg   1200

agtcctccag ggtccccatg ccctgagaga atttctaggg aagtcatctc acttggcctt   1260

ctgaaggtcc tccctaagag tctcctgaca aaagttactt attga   1305

<210> 14
<211> 434
<212> PRT
<213> Homo sapiens

<400> 14

Met Gly Arg Gly Leu Trp Glu Ala Trp Pro Pro Ala Gly Ser Ser Ala
1   5   10   15

Val Ala Lys Gly Asn Cys Arg Glu Glu Ala Glu Gly Ala Glu Asp Arg
            20   25   30

Gln Pro Ala Ser Arg Arg Gly Ala Gly Thr Thr Ala Ala Met Ala Ala
            35   40   45

Ser Gly Pro Gly Cys Arg Ser Trp Cys Leu Cys Pro Glu Val Pro Ser
    50   55   60

Ala Thr Phe Phe Thr Ala Leu Leu Ser Leu Leu Val Ser Gly Pro Arg
65   70   75   80

Leu Phe Leu Leu Gln Gln Pro Leu Ala Pro Ser Gly Leu Thr Leu Lys
            85   90   95

Ser Glu Ala Leu Arg Asn Trp Gln Val Tyr Arg Leu Val Thr Tyr Ile
            100   105   110

Phe Val Tyr Glu Asn Pro Ile Ser Leu Leu Cys Gly Ala Ile Ile Ile
        115   120   125

Trp Arg Phe Ala Gly Asn Phe Glu Arg Thr Val Gly Thr Val Arg His
    130   135   140

Cys Phe Phe Thr Val Ile Phe Ala Ile Phe Ser Ala Ile Ile Phe Leu
145   150   155   160

Ser Phe Glu Ala Val Ser Ser Leu Ser Lys Leu Gly Glu Val Glu Asp
            165   170   175

Ala Arg Gly Phe Thr Pro Val Ala Phe Ala Met Leu Gly Val Thr Thr
            180   185   190

Val Arg Ser Arg Met Arg Arg Ala Leu Val Phe Gly Met Val Val Pro
        195   200   205

Ser Val Leu Val Pro Trp Leu Leu Leu Gly Ala Ser Trp Leu Ile Pro
    210   215   220

Gln Thr Ser Phe Leu Ser Asn Val Cys Gly Leu Ser Ile Gly Leu Ala
225   230   235   240

Tyr Ala His Leu Leu Leu Phe His Arg Pro Leu Arg Ala Ser Gly Ala
            245   250   255

22

```
Glu Ala Arg Ser Asp Leu Pro Leu Gln Pro Asp Glu Glu Asp Ile Arg
            260             265             270

Val Gln Val Arg Leu Arg Val Phe Ser Arg Glu Glu Gly Ser Pro Glu
            275             280             285

Pro Glu Thr Glu Pro Gly Ala Trp Leu Leu Pro His Thr Glu Leu Pro
        290             295             300

Pro Ser Pro Val Pro Lys Pro Pro Cys Val Pro Asp Ala Ala Arg Gln
305             310             315             320

Trp Ser Glu Ala Gly Leu Leu Ala Ser Cys Thr Pro Gly His Met Pro
                325             330             335

Thr Leu Pro Pro Tyr Gln Pro Ala Ser Gly Leu Cys Tyr Val Gln Asn
            340             345             350

His Phe Gly Pro Asn Pro Thr Ser Ser Ser Val Tyr Pro Ala Ser Ala
        355             360             365

Gly Thr Ser Leu Gly Ile Gln Pro Pro Thr Pro Val Asn Ser Pro Gly
    370             375             380

Thr Val Tyr Ser Gly Ala Leu Gly His Gln Gly Leu Gln Ala Pro Arg
385             390             395             400

Ser Pro Pro Gly Ser Pro Cys Pro Glu Arg Ile Ser Arg Glu Val Ile
                405             410             415

Ser Leu Gly Leu Leu Lys Val Leu Pro Lys Ser Leu Leu Thr Lys Val
            420             425             430

Thr Tyr
```

```
<210>  15
<211>  1738
<212>  DNA
<213>  Homo sapiens

<400>  15
tgctgactaa tccgcgggcc gcggggaatg ggtggcgcgc cagccggtcg gcagggggtca      60

cgcggccggg tgtgcgcgga atggattcag ggttctcggg tcgcgcccgg gagctagagt     120

cctgactccc agccggggtc cccgaccggc ctttcggggt tgccgggcgc gctctgcaaa     180

ggggaaactg aggcccaagg aaattgagta tctctgcaaa gtcaccagct gagttcgaac     240
```

23

```
cagacagatc ctgggacccc ctactctatc cgccaccgga agaatctccg tccttgtctt      300

tccattctgc cctcccggct tcccagtaag cacacagccc caggaccacc cttgaccgtc      360

ctcaaccaag cgatgcatgc caggggcccc catggccaac tgtccccagc actgcctctg      420

gcctcctcag tcctgatgct gctgatgagc accctgtggc tggtggggggc cggccccggc      480

ctggtcctgg ccccggagct gttgctggac ccctggcagg tgcaccggct gctgacccat      540

gccctgggcc acacggccct gccaggcctg ctcctgagcc tgctgctcct gcccactgtg      600

ggctggcagc aggagtgcca cctgggcacg ctgagattcc tgcatgcctc agccctgctc      660

gccctggctt ctgggctgct ggcagtgctg ctggcaggcc ttgggctgtc cagtgcagcc      720

ggcagctgtg gatacatgcc tgtccacctg gccatgctgg ctggggaggg acaccgccct      780

agacggcccc gtggggcact gccaccgtgg ctgtcgccgt ggctgctgct tgccctgacc      840

ccactgctca gctctgagcc acccttcctg cagctccttt gcggcctcct gccggcctg      900

gcctatgcag ctggggcctt ccggtggctg gaaccctcag agcgacggct gcaggtgctg      960

caggagggcg tcttgtgcag gaccttggcg gggtgctggc ccctgaggct ccttgccacc     1020

ccgggtagcc tggcggagct gcctgtcacc catcctgccg gagtgaggcc tcccatccct     1080

ggaccgcctt atgtggcctc ccctgacctc tggtcccact gggaagactc agccctgccc     1140

ccaccaagcc tgaggcctgt gcagcccacc tgggagggct cctcagaggc aggcctggac     1200

tgggctgggg ccagcttctc cccagggact ccgatgtggg cggccttgga tgagcagatg     1260

ctgcaggagg gcatccaggc ctcgcttctt gacgggccag cccaggaacc ccagagcgca     1320

ccatggctgt ccaagtcctc tgtctcctct ctgcggctgc agcagctgga gcgcatgggc     1380

ttccctacgg agcaggcggt ggtggcactg gcagccacag gccgtgtgga gggtgccgtg     1440

tcactgttgg ttggaggaca agtgggcact gagaccctgg tgacccatgg aaagggtggg     1500

cctgcccact ccgagggtcc tgggcctccc tagcccaggc agagagtggg gcacaggcag     1560

gcccttgggt gctaagggct gggctgcatg tgggtagccc gagctcctac tctgtctaaa     1620

gagggccaca gtggggagca ggggcacctc tggaggcagg agaggccccc cagcatgctg     1680

ccctagtacg tgtttagaat aaaaaccagt ttgttttttca acctggacct ccttggag      1738
```

```
<210>   16
<211>   386
<212>   PRT
<213>   Homo sapiens

<400>   16

Met His Ala Arg Gly Pro His Gly Gln Leu Ser Pro Ala Leu Pro Leu
1               5                   10                  15
```

24

```
Ala Ser Ser Val Leu Met Leu Leu Met Ser Thr Leu Trp Leu Val Gly
            20                  25              30


Ala Gly Pro Gly Leu Val Leu Ala Pro Glu Leu Leu Leu Asp Pro Trp
            35                  40              45


Gln Val His Arg Leu Leu Thr His Ala Leu Gly His Thr Ala Leu Pro
    50                  55                  60


Gly Leu Leu Leu Ser Leu Leu Leu Pro Thr Val Gly Trp Gln Gln
65                  70                  75                  80


Glu Cys His Leu Gly Thr Leu Arg Phe Leu His Ala Ser Ala Leu Leu
                85                  90                  95


Ala Leu Ala Ser Gly Leu Leu Ala Val Leu Leu Ala Gly Leu Gly Leu
            100                 105                 110


Ser Ser Ala Ala Gly Ser Cys Gly Tyr Met Pro Val His Leu Ala Met
        115                 120                 125


Leu Ala Gly Glu Gly His Arg Pro Arg Arg Pro Arg Gly Ala Leu Pro
    130                 135                 140


Pro Trp Leu Ser Pro Trp Leu Leu Leu Ala Leu Thr Pro Leu Leu Ser
145                 150                 155                 160


Ser Glu Pro Pro Phe Leu Gln Leu Leu Cys Gly Leu Leu Ala Gly Leu
                165                 170                 175


Ala Tyr Ala Ala Gly Ala Phe Arg Trp Leu Glu Pro Ser Glu Arg Arg
            180                 185                 190


Leu Gln Val Leu Gln Glu Gly Val Leu Cys Arg Thr Leu Ala Gly Cys
        195                 200                 205


Trp Pro Leu Arg Leu Leu Ala Thr Pro Gly Ser Leu Ala Glu Leu Pro
210                 215                 220


Val Thr His Pro Ala Gly Val Arg Pro Pro Ile Pro Gly Pro Pro Tyr
225                 230                 235                 240


Val Ala Ser Pro Asp Leu Trp Ser His Trp Glu Asp Ser Ala Leu Pro
                245                 250                 255


Pro Pro Ser Leu Arg Pro Val Gln Pro Thr Trp Glu Gly Ser Ser Glu
            260                 265                 270
```

25

Ala Gly Leu Asp Trp Ala Gly Ala Ser Phe Ser Pro Gly Thr Pro Met
        275   280   285

Trp Ala Ala Leu Asp Glu Gln Met Leu Gln Glu Gly Ile Gln Ala Ser
    290   295   300

Leu Leu Asp Gly Pro Ala Gln Glu Pro Gln Ser Ala Pro Trp Leu Ser
305   310   315   320

Lys Ser Ser Val Ser Ser Leu Arg Leu Gln Gln Leu Glu Arg Met Gly
                325   330   335

Phe Pro Thr Glu Gln Ala Val Val Ala Leu Ala Ala Thr Gly Arg Val
            340   345   350

Glu Gly Ala Val Ser Leu Leu Val Gly Gly Gln Val Gly Thr Glu Thr
        355   360   365

Leu Val Thr His Gly Lys Gly Gly Pro Ala His Ser Glu Gly Pro Gly
    370   375   380

Pro Pro
385

<210> 17
<211> 19
<212> PRT
<213> Homo sapiens

<400> 17

Met Gly Arg Val Glu Asp Gly Gly Thr Thr Glu Glu Leu Glu Asp Trp
1   5   10   15

Asp Pro Gly

<210> 18
<211> 1559
<212> DNA
<213> Homo sapiens

<400> 18
cttggacctt ggccctcgct ttccaggatg ggtagggtgg aagacggggg aacaactgag   60

gagctggagg actgggaccc aggcaccagt gccctgccag ctcctgggat caagcagggt   120

cccagggaac agacaggcac ggggcccctg tcccaaaagt gctgggagcc tgagcctgat   180

gctcccagcc agcctggccc agccctttgg tccaggggtc gggcccgcac tcaggccttg   240

26

gctggcggct cctcactgca gcagctggac cccgagaaca caggcttcat cggtgcggac   300

accttcactg gcctggtgca cagccatgag ctgcccctgg acccggccaa gctggacatg   360

ctggtggccc tggctcagag caacgagcag ggccaggtct gctaccagga gctggtggac   420

ctgatcagca gcaagcgctc cagcagtttc aagcgggcca ttgctaacgg acagcgggca   480

ctgccccggg acgggccgct ggatgagcca ggcctaggtg tctacaagcg gtttgtgcgt   540

tacgtggcct acgagatcct gccttgtgag gtggaccgcc gctggtactt ctaccgtcac   600

cgcagctgcc cacccccgt gttcatggcc tcggtcactc ttgcccagat catcgtgttc   660

ctgtgttacg gggcccgcct caacaagtgg gtgctgcaga cctaccaccc cgagtacatg   720

aagagccccc ttgtgtacca ccccgggcac cgtgcccgcg cctggcgctt cctcacctac   780

atgttcatgc acgttgggct ggagcagctg gggttcaacg ccctcctgca gctgatgatc   840

ggggtgcccc tggagatggt gcacggcctg ctccgcatca gcctgctcta cctggcaggc   900

gtgctggcag gctccctaac cgtctccatc accgacatgc gggccccggt ggtgggaggc   960

tccggcgggg tctacgccct gtgctcggca cacctggcca acgttgtcat gaactgggct   1020

gggatgagat gtccctacaa gttgctgagg atggtgctgg ccttggtgtg catgagctcc   1080

gaggtgggcc gggccgtgtg gctgcgcttc tccccgccgc tgcccgcctc gggcccacag   1140

cccagcttca tggcgcacct ggcaggcgcg gtggtggggg tgagcatggg cctgaccatc   1200

ctgcggagct acgaggagcg cctgcgggac cagtgcggct ggtgggtggt gctgctggcc   1260

tacggcacct tcctgctctt cgccgtcttc tggaacgtct tcgcctacga cctgctgggc   1320

gcccacatcc ccccaccgcc ctgaccggct acctgaggct gcacaggcca gggctcgggc   1380

atgtggtggc cgcccaccag gggccttcac gtctgccctt tgtgaacgga cgtctcaggg   1440

ctgctgtgcc ccttgggtgt gggtggcctc aaaggaggcc ctgtcccagc cacccacccc   1500

ccactcccag gacttgcggt ctgagccttt ttggataatt aataaatatt ttacacagc   1559

<210>   19
<211>   593
<212>   DNA
<213>   Homo sapiens

<400>   19
atttactatg ctgtgtggaa gcctcagaaa cagtggatca cgttggacac aggcatcttg   60

gagagtccct ttatctacag tcctgagaag agggaggaag cctggaggtt tatctcatac   120

atgctggtac atgctggagt tcagcacatc ttggggaatc tttgtatgca gcttgttttg   180

ggtattccct tggaaatggt ccacaaaggc ctccgtgtgg ggctggtgta cctggcagga   240

gtgattgcag ggtcccttgc cagctccatc tttgacccac tcagatatct tgtgggagct   300

27

```
tcaggaggag tctatgctct gatgggaggc tattttatga atgttctggt gaattttcaa     360

gaaatgattc ctgcctttgg aattttcaga ctgctgatca tcatcctgat aattgtgttg     420

gacatgggat ttgctctcta tagaaggttc tttgttcctg aagatgggtc tccggtgtct     480

tttgcagctc acattgcagg tggatttgct ggaatgtcca ttggctacac ggtgtttagc     540

tgctttgata aagcactgct gaaagatcca aggttatgga tagcaattgc tgc           593
```

```
<210>   20
<211>   708
<212>   DNA
<213>   Homo sapiens

<400>   20
aaggaggagc gcctggcaga gcgggatctt cagggcagcg ggcaaccct tggcccagga      60

agcctggaac gcaaggaccg gagggcgtgg gctgggacgc ccctacgttg gtctttcagg     120

gaaaggcctt ggaaagcagt cgttgcgcca gacagcccag ggaagagcgg cagcctgagg     180

acctagggcc acctgctgtt ccctgggatt catgtccttc tggggaggag ggaggaccca     240

ggacaatggc tgctgttcat gatctggaga tggagagcat gaatctgaat atggggagag     300

agatgaaaga agagctggag gaagaggaga aaatgagaga ggatggggga ggtaaagatc     360

gggccaagag taaaaaggtc cacaggattg tctcaaaatg gatgctgccc gaaaagtccc     420

gaggaacata cttggagaga gctaactgct tcccgcctcc cgtgttcatc atctccatca     480

gcctggccga gctggcagtg tttatttact atgctgtgtg gaagcctcag aaacagtgga     540

tcacgttgga cacaggcatc ttggagagtc cctttatcta cagtcctgag aagagggagg     600

aagcctggag gtttatctca tacatgctgg tacatgctgg gtaagcaatg atagttaagc     660

cctggtatca gaggtgatta tatcattgta acctcttaaa acaaaaaa                  708
```

```
<210>   21
<211>   616
<212>   DNA
<213>   Homo sapiens


<220>
<221>   misc_feature
<222>   (498)..(498)
<223>   n is a, c, g, or t

<400>   21
cttcttccct ttctacatta agaagcccct tccttaaatc ctaaggtcct tttataggca      60

atctttgcaa ctgatgagtt taatgctgtt ttgtcctcta tataaaattg ttagcctttg     120

ctgagacttc tctgtttctt catagaggct tccctttttt tttttatttc ctccagatgg     180

ctctttttat taattgactt acaatagggg caggtcagtt tgctggagat aggaaaatgt     240
```

28

```
tgaaaaacac agcaaataag acacaagcta aatatgcagc aattgctatc caaaaccttg      300

gatctttcat cagtgcttta tcaaagcagc taaacaccgt gtagccaatg gacattccag      360

caaatccacc tgcaatgtga gctgcaaaag acaccggaga cccatcttca ggaacaaaga      420

accttctata gagagcaaat cccatgtcca acacaattat caggatgatg atcagcagtc      480

tgaaaattcc aaaggcanga atcatttctt gaaaattcac cagaacattc ataaaatagc      540

ctcccatcag agcatagact cctcctgaag ctcccacaag atatctgagt gggtcaaaga      600

tggagctggc aaggga                                                     616
```


<210> 22
<211> 354
<212> DNA
<213> Homo sapiens

<400> 22
```
tggatcacgt tggacacagg catttggaga gtcccttttat ctacagtcct gagaagaggg       60

aggaagcctg gaggtttatc tcacacatgc tggtacatgc tggagttcag cacatcttgg      120

ggaatctttg tatgcagctt gttttgggta ttcccttgga aatggtccac aaaggcctcc      180

gtgtggggct ggtgtacctg gcaggagtga ttgcagggtc ccttgccagc tccatctttg      240

acccactcag atatcttgtg ggagcttcag gaggagtcta tgctctgatg ggaggctatt      300

ttatgaatgt tctggtgaat tttcaagaaa tgattcctgc ctttggaatt ttca           354
```


<210> 23
<211> 356
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (356)..(356)
<223> n is a, c, g, or t

<400> 23
```
tttaaaacaa ttacagagaa aactttatttt tgggccattt aggaggttta gatcattttg       60

atcatcttca gctgtcttct cttcacatac aggaaaggcc ttggaaagca gtcgttgcgc      120

cagacagccc agggaagagc ggcacgctga ggacctaggg ccacctgctg ttccctggga      180

ttcatgtcct tctggggagg agggaggacc caggacaatg gctgctgttc atgatctgga      240

gatggagagc atgaatctga atatggggag agagatgaaa gaagagctgg aggaagagga      300

gaaaatgaga gaggatgggg gaggtaaaga tcgggccaag agtaaaaagg tccacn         356
```


<210> 24
<211> 524
<212> DNA
```

29

<213>  Homo sapiens

<400>  24
gggtctccgg tgtcttttgc agctcacatt gcaggtggat ttgctggaat gtccattggc      60

tacacggtgt ttagctgctt tgataaagca ctgatgaaag atccaaggtt ttggatagca     120

attgctgcat atttagcttg tgtcttattt gctgtgtttt tcaacatttt cctatctcca     180

gcaaactgac ctgcccctat tgtaagtcaa ttaataaaaa gagccatctg gaggaaataa     240

aaaaaaaagg aagactctat gaagaaacag agaagtctca gcaaaggcta acaattttat     300

atagaggaca aaacagcatt aaactcatca gttgcaaaga ttgcctataa aaggacctta     360

ggatttaagg aaggggcttc ttaatgtaga aagggaagaa gaagagagaa aagaagggta     420

gtaaaaacta gagattgggg ccaggcgcag tggctcacgc ctgtaatccc agcactttgg     480

gaggctgagg cgggtgaaat cacctgaggt catgagttca agac                      524


<210>  25
<211>  428
<212>  DNA
<213>  Homo sapiens

<400>  25
ttttttttcc ctttctacat taagaagccc cttccttaaa tcctaaggtc cttttatagg      60

caatctttgc aactgatgag tttaatgctg ttttgtcctc tatataaaat tgttagcctt     120

ttctgagact tctctgtttc ttcatagagt cttccttttt tttattattt cctccagatg     180

gctcttttta ttaattgact tacaataggg gcaggtcagt ttgctggaga taggaaaatg     240

ttgaaaaaca cagcaaataa gacacaagct aaatatgcag caattgctat ccaaaacctt     300

ggatctttca gcagtgcttt atcaaagcag ctaaacaccg tgtagccaat ggacattcca     360

gcaaatccac ctgcaatgtg agctgcaaaa gacaccttgg gagaaaagga gggaaaatgg     420

aaataagt                                                              428


<210>  26
<211>  271
<212>  DNA
<213>  Homo sapiens

<400>  26
ctcgaaactg tgaaaacaca gcaaataaga cacaagctaa atatgcagca attgctatcc      60

aaaaccttgg atctttcagc agtgctttat caaagcagct aaacaccgtg tagccaatgg     120

acattccagc aaatccacct gcaatgtgag ctgcaaaaga cacctgggag aaaaggaggg     180

aaaatggaaa tcagtgaaga caatgctaat tgtgtatttc agttgcatct ctcttattca     240

aacacaaata ccgtcacagt taaacaagcg t                                    271

```
<210>   27
<211>   262
<212>   DNA
<213>   Homo sapiens

<400>   27
tgaaatacac agcaaataga cacaagctaa atatgcagca attgctatcc aaaagcttgg      60

atctttcagc agtgctttat caaagcagct aaacaccgtg tagcgcatgg acattccagc     120

aaatccacct gcaatgtgag ctgcaaaaga caccttggga gaaaaggagg gaaaatggaa     180

ataagtgaag acaatgctaa ttgtgtattt cagttgcatc tctcttattc aaacacaaat     240

accgtcacag ttaaacaagc gt                                              262


<210>   28
<211>   120
<212>   DNA
<213>   Homo sapiens

<400>   28
gctcggtccg aaactgttga aaaacacagc aaataagaca caacgctaaa tatgcagcaa      60

ttgctatcca aaaccttgga tctttcagca gtgctttatc aaagcagcta aacaccgtgt     120


<210>   29
<211>   894
<212>   DNA
<213>   Homo sapiens

<400>   29
ctcatccacg ttacttgaag aactccctgg tttaccaccc acagctgcga gcacaggttt      60

ggcgctacct gacatacatc ttcatgcatg cagggtatag aacacctgga ctcaatgtgg     120

tgctgcagct gtggtggggg tgcccctgga gatggtgcat ggagccaccc gaattgggct     180

tgtctacgtg gccggtgttg tggcagggtc cttggcagtg tctagtggct gacatgaccg     240

ctccagtcgt gggctcttct ggaggggtgt atgctctcgt ctctgcccat ctggccaaca     300

ttgtcatgaa ctggtcaggc atgaagtgcc agttcaagct gctgcggatg gctgtggccc     360

ttatctgtat gagcatggag cttgggcggg ccgtgtggct ccgcttccac ccgtcggcta     420

tcccccgtgc cctcacccaa gctttgtggc gcacttgggt ggcgtggccg tgtgcatcac     480

cctggccgtg gtggtcctga ggaactacga gcagaggctc caggaccagt cactgtggtg     540

gattttgtgt gccatgtaca ccgtcttcgg gctgttcgct gtctctggaa catctttggc     600

tacacctgtt ggacttaaag ctgcgctccc ccctgaggct tgaggcccca gtccgcgagg     660

ggagggaaag caggacccca gggagcgcct ggaaggttct tcctcatcac aggttcagta     720

agcgggcaaa cacgcacaaa cactgcgcct gtattgtgtt atttggccac gggggcacct     780

ctgcgaaggg ctgcgggcga aacacaagac aaacggacta acacaagaac aagggggcgaa     840
```

tcatcgcgcg ggagcacttg gaaagcaggt taaaacggga cacaaggaca cttt   894

<210> 30
<211> 344
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (19)..(19)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (49)..(49)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (92)..(92)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (130)..(130)
<223> n is a, c, g, or t

<400> 30
cctccgccag atgcctttna gaaaagggtc tccactgtgt gtccaaatnt aaacacaaac   60

attacagtgg cccagagtct tctgtccttg tntgcccggc ctagctgagc tgatgatgag   120

aagaaacctn gcaggcgctc cctgtgggtg ctgcttttcc ctcccctccc cgaccttggg   180

cctccagccc tcaggggga ggcggcagct ttaagtccag cagggtgtag gcaaagatgt   240

tccagaagac agcgaacagc acgaagacgg tgtacatggc cacaaaaatc caccacagtg   300

actggtcctg gagcctctgc tcgtagttcc tcaggaccac cacg   344

<210> 31
<211> 1140
<212> DNA
<213> Homo sapiens

<400> 31
atggcgtggc gaggctgggc gcagagaggc tggggctgcg gccaggcgtg gggtgcgtcg   60

gtgggcggcc gcagctgcga ggagctcact gcggtcctaa ccccgccgca gctcctcgga   120

cgcaggttta acttctttat tcaacaaaaa tgcggattca gaaaagcacc caggaaggtt   180

gaacctcgaa gatcagaccc agggacaagt ggtgaagcat acaagagaag tgctttgatt   240

cctcctgtgg aagaaacagt cttttatcct tctccctatc ctataaggag tctcataaaa   300

cctttatttt ttactgttgg gtttacaggc tgtgcatttg gatcagctgc tatttggcaa   360

tatgaatcac tgaaatccag ggtccagagt tattttgatg gtataaaagc tgattggttg   420

```
gatagcataa gaccacaaaa agaaggagac ttcagaaagg agattaacaa gtggtggaat        480

aacctaagtg atggccagcg gactgtgaca ggtattatag ctgcaaatgt ccttgtattc        540

tgtttatgga gagtaccttc tctgcagcgg acaatgatca gatatttcac atcgaatcca        600

gcctcaaagg tcctttgttc tccaatgttg ctgtcaacat tcagtcactt ctccttattt        660

cacatggcag caaatatgta tgttttgtgg agcttctctt ccagcatagt gaacattctg        720

ggtcaagagc agttcatggc agtgtaccta tctgcaggtg ttatttccaa ttttgtcagt        780

tacctgggta aagttgccac aggaagatat ggaccatcac ttggtgcatc tggtgccatc        840

atgacagtcc tcgcagctgt ctgcactaag atcccagaag ggaggcttgc cattattttc        900

cttccgatgt tcacgttcac agcagggaat gccctgaaag ccattatcgc catggataca        960

gcaggaatga tcctgggatg gaaatttttt gatcatgcgg cacatcttgg gggagctctt       1020

tttggaatat ggtatgttac ttacggtcat gaactgattt ggaagaacag ggagccgcta       1080

gtgaaaatct ggcatgaaat aaggactaat ggccccaaaa aaggaggtgg ctctaagtaa       1140
```

```
<210>   32
<211>   2964
<212>   DNA
<213>   Homo sapiens
```

```
<400>   32
gactcggcgc gggcgccctc ccggccagcg gcggcagccc ctcctccccg gcgccctcag         60

gacccccag agacccccgg cggcggcagc ctgccttgct ctgccaggaa ccatgagtga        120

ggcccgcagg gacagcacga gcagcctgca gcgcaagaag ccaccctggc taaagctgga        180

cattccctct gcggtgcccc tgacggcaga agagcccagc ttcctgcagc ccctgaggcg        240

acaggctttc ctgaggagtg tgagtatgcc agccgagaca gcccacatct cttcaccCca        300

ccatgagctc cggcggccgg tgctgcaacg ccagacgtcc atcacacaga ccatccgcag        360

ggggaccgcc gactggtttg gagtgagcaa ggacagtgac agcacccaga aatggcagcg        420

caagagcatc cgtcactgca gccagcgcta cgggaagctg aagccccagg tcctccggga        480

gctggacctg cccagccagg acaacgtgtc gctgaccagc accgagacgc accccccact        540

ctacgtgggg ccatgccagc tgggcatgca agagatcata gaccccctgg cccgtggccg        600

tgccttccgt gtggcagatg acactgcgga aggcctgagt gccccacaca ctcccgtcac        660

gccgggtgct gcctccctct gctccttctc cagctcccgc tcaggtttcc accggctccc        720

gcggcggcgc aagcgagagt cggtggccaa gatgagcttc cgggcggccg cagcgctgat        780

gaaaggccgc tccgttaggg atggcacctt cgccgggca cagcgtcgaa gcttcactcc        840

agctagcttt ctggaggagg acacaactga tttcCccgat gagctggaca catccttctt        900
```

```
tgcccgggaa ggtatcctcc atgaagagct gtccacatac ccggatgaag ttttcgagtc    960

cccatcggag gcagcgctaa aggactggga gaaggcaccg gagcaggcgg acctcaccgg   1020

cggggccctg gaccgcagcg agcttgagcg cagccacctg atgctgccct tggagcgagg   1080

ctggcggaag cagaaggagg gcgccgcagc cccgcagccc aaggtgcggc tccgacagga   1140

ggtggtgagc accgcggggc cgcgacgggg ccagcgtatc gcggtgccgg tgcgcaagct   1200

cttcgcccgg gagaagcggc cgtatgggct gggcatggtg ggacggctca ccaaccgcac   1260

ctaccgcaag cgcatcgaca gcttcgtcaa gcgccagatc gaggacatgg acgaccacag   1320

gcccttcttc acctactggc ttaccttcgt gcactcgctc gtcgccatcc tagccgtgtg   1380

catctatggc atcgcgcccg tgggcttctc gcagcatgag acggtggact cggtgctgcg   1440

gaaccgcggg gtctacgaga cgtcaagta cgtgcagcag gagaacttct ggatcgggcc   1500

cagctcggag gccctcatcc acctgggcgc caagttttcg ccctgcatgc gccaggaccc   1560

gcaggtgcac agcttcattc gctcggcgcg cgagcgcgag aagcactccg cctgctgcgt   1620

gcgcaacgac aggtcgggct gcgtgcagac ctcggaggag gagtgctcgt ccacgctggc   1680

agtgtgggtg aagtggccca tccatcccag cgccccagag cttgcgggcc acaagagaca   1740

gtttggctct gtctgccacc aggatcccag ggtgtgtgat gagccctcct ccgaagaccc   1800

tcatgagtgg ccagaagaca tcaccaagtg gccgatctgc accaaaaaca gcgctgggaa   1860

ccacaccaac catccccaca tggactgtgt catcacagga cggccctgct gcattggcac   1920

caagggcagg tgtgagatca cctcccggga gtactgtgac ttcatgaggg gctacttcca   1980

tgaggaggcc acgctctgct ctcaggtgca ctgcatggat gatgtgtgtg ggctcctgcc   2040

ttttctcaac cccgaggtgc ctgaccagtt ctaccgcctg tggctatccc tcttcctgca   2100

cgccgggatc ttgcactgcc tggtgtccat ctgcttccag atgactgtcc tgcgggacct   2160

ggagaagctg gcaggctggc accgcatagc catcatctac ctgctgagtg gtgtcaccgg   2220

caacctggcc agtgccatct tcctgccata ccgagcagag gtgggtcctg ctggctccca   2280

gttcggcatc ctggcctgcc tcttcgtgga gctcttccag agctggcaga tcctggcgcg   2340

gccctggcgt gccttcttca gctgctggc tgtggtgctc ttcctcttca cctttgggct   2400

gctgccgtgg attgacaact ttgcccacat ctcggggttc atcagtggcc tcttcctctc   2460

cttcgccttc ttgccctaca tcagctttgg caagttcgac ctgtaccgga aacgctgcca   2520

gatcatcatc tttcaggtgg tcttcctggg cctcctggct ggcctggtgg tcctcttcta   2580

cgtctatcct gtccgctgtg agtggtgtga gttcctcacc tgcatcccct tcactgacaa   2640

gttctgtgag aagtacgaac tggacgctca gctccactga gctggctgcg ggctccagcg   2700

gccgtgtgct ccagcaggcc agagccagac acgacctccc tgagcctcac aggcttacag   2760
```

34

```
gagtcacctg ctccatgtgg ggactggcct gtttcctgaa cacagacctc tttcttgtgc      2820

cttgttcact tctgttgaac ccctcgtact gccgggcatt tattatacta cttcctgtca      2880

taaccttcta acttgtttct tgacgaccac ctcatgtggc caataaatga actgggagcg      2940

tttttaaaaaa aaaaaaaaaa aaaa                                            2964
```

<210> 33
<211> 2643
<212> DNA
<213> Homo sapiens

<400> 33
```
gctactcccca cctgccacgc cgcaagagaa tgtctgtggc ccacatgagc ttgcaagctg       60

ccgctgccct cctcaagggg cgctcggtgc tggatgccac cggacagcgg tgccgggtgg      120

tcaagcgcag ctttgccttc ccgagcttcc tggaggagga tgtggtcgat ggggcagaca      180

cgtttgactc ctcctttttt agtaaggaag aaatgagctc catgcctgat gatgtctttg      240

agtccccccc actctctgcc agctacttcc gagggatccc acactcagcc tcccctgtct      300

cccccgatgg ggtgcaaatc cctctgaagg agtatggccg agccccagtc cccgggcccc      360

ggcgcggcaa gcgcatcgcc tccaaggtga agcactttgc ctttgatcgg aagaagcggc      420

actacggcct cggcgtggtg ggcaactggc tgaaccgcag ctaccgccgc agcatcagca      480

gcactgtgca gcggcagctg gagagcttcg acagccaccg gccctacttc acctactggc      540

tgaccttcgt ccatgtcatc atcacgctgc tggtgatttg cacgtatggc atcgcacccg      600

tgggctttgc ccagcacgtc accacccagc tggtgctgcg gaacaaaggt gtgtacgaga      660

gcgtgaagta catccagcag gagaacttct gggttggccc cagctcgatt gacctgatcc      720

acctgggggc caagttctca ccctgcatcc ggaaggacgg gcagatcgag cagctggtgc      780

tgcgcgagcg agacctggag cgggactcag gctgctgtgt ccagaatgac cactccggat      840

gcatccagac ccagcggaag gactgctcgg agctttggc cacttttgtc aagtggcagg      900

atgacactgg gccccccatg gacaagtctg atctgggcca gaagcggact tcggggggctg      960

tctgccacca ggaccccagg acctgcgagg agccagcctc cagcggtgcc cacatctggc     1020

ccgatgacat cactaagtgg ccgatctgca cagagcaggc caggagcaac cacacaggct     1080

tcctgcacat ggactgcgag atcaagggcc gcccctgctg catcggcacc aagggcagct     1140

gtgagatcac cacccgggaa tactgtgagt tcatgcacgg ctatttccat gaggaagcaa     1200

cactctgctc ccaggtgcac tgcttggaca aggtgtgtgg gctgctgccc ttcctcaacc     1260

ctgaggtccc agatcagttc tacaggctct ggctgtctct cttcctacat gctggcgtgg     1320

tgcactgcct cgtgtctgtg gtctttcaaa tgaccatcct gagggacctg gagaagctgg     1380

ccggctggca ccgtatcgcc atcatcttca tcctcagtgg catcacaggc aacctcgcca     1440
```

```
gtgccatctt tctcccatac cgggcagagg tgggcccggc cggctcacag ttcggcctcc    1500

tcgcctgcct cttcgtggag ctcttccaga gctggccgct gctggagagg ccctggaagg    1560

ccttcctcaa cctctcggcc atcgtgctct tcctgttcat ctgtggcctc ttgccctgga    1620

tcgacaacat cgcccacatc ttcggcttcc tcagtggcct gctgctggcc ttcgccttcc    1680

tgccctacat caccttcggc accagcgaca agtaccgcaa gcgggcactc atcctggtgt    1740

cactgctggc ctttgccggc ctcttcgccg ccctcgtgct gtggctgtac atctacccca    1800

ttaactggcc ctggatcgag cacctcacct gcttcccctt caccagccgc ttctgcgaga    1860

agtatgagct ggaccaggtg ctgcactgac cgctgggcca cacggctgcc cctcagccct    1920

gctggaacag ggtctgcctg cgagggctgc cctctgcaga gcgctctctg tgtgccagag    1980

agccagagac ccaagacagg gcccgggctc tggacctggg tgcccccctg ccaggcgagg    2040

ctgactccgc gtgagatggt tggttaaggc ggggtttttc tggggcgtga ggcctgtgag    2100

atcctgaccc aagctcaggc acacccaagg cacctgcctc tctgagtctt gggtctcagt    2160

tcctaatatc ccgctccttg ctgagaccat ctcctggggc agggtccttt tcttcccagg    2220

tcctcagcgc tgcctctgct ggtgccttct cccccactac tactggagcg tgcccttgct    2280

ggggacgtgg ctgtgccctc agttgccccc agggctgggt gcccaccatg ccccttcctc    2340

tttctcctcc tacctctgcc ctgtgagccc atccataagg ctctcagatg ggacattgta    2400

ggaaaggctt tggccatggt ctgggggcag agaacaaggg gggagacaca agtagacctc    2460

aggtagaacg acaccgggcg gagccacccc agggcctgct cccagggagt gctcgaggcg    2520

catcaggccc gttttttacc agtttatatc acggtcttca ttttaaaag taacgctaac     2580

tttgtacgga cgatgtctca tggattaaat aatattcttt atggcaaaaa aaaaaaaaa     2640

aaa                                                                   2643
```

```
<210>    34
<211>    1305
<212>    DNA
<213>    Homo sapiens

<400>    34
atggggcggg gcctctggga ggcgtggcct ccggccggct cctctgctgt tgccaaggga      60

aactgccgcg aggaggcgga aggagcagag gaccggcagc cggcgtcgag gcggggcgcg     120

ggaacgacgg cggccatggc ggcctcgggg cccgggtgtc gcagctggtg cttgtgtccc     180

gaggtgccat ccgccacctt cttcactgcg ctgctctcgc tgctggtttc cgggcctcgc     240

ctgttcctgc tgcagcagcc cctggcgccc tcgggcctca cgctgaagtc cgaggccctt     300

cgcaactggc aagtttacag gctggtaacc tacatctttg tctacgagaa tcccatctcc     360
```

```
ctgctctgcg gcgctatcat catctggcgc tttgctggca atttcgagag aaccgtgggc      420

accgtccgcc actgcttctt caccgtgatc ttcgccatct tctccgctat catcttcctg      480

tcattcgagg ctgtgtcatc actgtcaaag ctgggggaag tggaggatgc cagaggtttc      540

accccagtgg cctttgccat gctgggagtc accaccgtcc gttctcggat gaggcgggcc      600

ctggtgtttg gcatggttgt gccctcagtc ctggttccgt ggctcctgct gggtgcctcg      660

tggctcattc cccagacctc tttcctcagt aatgtctgcg ggctgtccat cgggctggcc      720

tatgctcacc tactgctatt ccatcgacct ctcagagcga gtggcgctga agctcgatca      780

gaccttcccc ttcagcctga tgaggaggat atccgtgttc aagtacgtct cagggtcttc      840

agccgagagg agggcagccc agagccggaa actgaacccg gtgcctggct cctaccccac      900

acagagctgc caccctcacc tgtccccaag ccaccctgtg tcccagacgc agcacgccag      960

tggtcagaag ctggcctcct ggcctcctgc acccccgggc acatgcccac cttgcctccg     1020

taccagcctg cctccggcct gtgctatgtg cagaaccact ttggtccaaa ccccacctcc     1080

tccagtgtct acccagcttc tgcgggcacc tccctgggca tccagccccc cacgcctgtg     1140

aacagccctg gcacggtgta ttctggggcc ttgggacacc aggggctgca ggctccaagg     1200

agtcctccag ggtccccatg ccctgagaga atttctaggg aagtcatctc acttggcctt     1260

ctgaaggtcc tccctaagag tctcctgaca aaagttactt attga                     1305
```

```
<210>   35
<211>   1738
<212>   DNA
<213>   Homo sapiens

<400>   35
tgctgactaa tccgcgggcc gcggggaatg ggtggcgcgc cagccggtcg gcaggggtca       60

cgcggccggg tgtgcgcgga atggattcag ggttctcggg tcgcgcccgg gagctagagt      120

cctgactccc agccggggtc cccgaccggc ctttcggggt tgccgggcgc gctctgcaaa      180

ggggaaactg aggcccaagg aaattgagta tctctgcaaa gtcaccagct gagttcgaac      240

cagacagatc ctgggacccc ctactctatc cgccaccgga agaatctccg tccttgtctt      300

tccattctgc cctcccggct tcccagtaag cacacagccc caggaccacc cttgaccgtc      360

ctcaaccaag cgatgcatgc caggggcccc catggccaac tgtccccagc actgcctctg      420

gcctcctcag tcctgatgct gctgatgagc accctgtggc tggtgggggc cggccccggc      480

ctggtcctgg ccccggagct gttgctggac ccctggcagg tgcaccggct gctgacccat      540

gccctgggcc acacggccct gccaggcctg ctcctgagcc tgctgctcct gcccactgtg      600

ggctggcagc aggagtgcca cctgggcacg ctgagattcc tgcatgcctc agccctgctc      660

gccctggctt ctgggctgct ggcagtgctg ctggcaggcc ttgggctgtc cagtgcagcc      720
```

                                      37

```
ggcagctgtg gatacatgcc tgtccacctg gccatgctgg ctggggaggg acaccgccct      780

agacggcccc gtggggcact gccaccgtgg ctgtcgccgt ggctgctgct tgccctgacc      840

ccactgctca gctctgagcc acccttcctg cagctccttt gcggcctcct tgccggcctg      900

gcctatgcag ctggggcctt ccggtggctg gaaccctcag agcgacggct gcaggtgctg      960

caggagggcg tcttgtgcag gaccttggcg gggtgctggc ccctgaggct ccttgccacc     1020

ccgggtagcc tggcggagct gcctgtcacc catcctgccg gagtgaggcc tcccatccct     1080

ggaccgcctt atgtggcctc ccctgacctc tggtcccact gggaagactc agccctgccc     1140

ccaccaagcc tgaggcctgt gcagcccacc tgggagggct cctcagaggc aggcctggac     1200

tgggctgggg ccagcttctc cccagggact ccgatgtggg cggccttgga tgagcagatg     1260

ctgcaggagg gcatccaggc ctcgcttctt gacgggccag cccaggaacc ccagagcgca     1320

ccatggctgt ccaagtcctc tgtctcctct ctgcggctgc agcagctgga gcgcatgggc     1380

ttccctacgg agcaggcggt ggtggcactg gcagccacag gccgtgtgga gggtgccgtg     1440

tcactgttgg ttggaggaca agtgggcact gagaccctgg tgacccatgg aaagggtggg     1500

cctgcccact ccgagggtcc tgggcctccc tagcccaggc agagagtggg gcacaggcag     1560

gcccttgggt gctaagggct gggctgcatg tgggtagccc gagctcctac tctgtctaaa     1620

gagggccaca gtggggagca ggggcacctc tggaggcagg agaggccccc cagcatgctg     1680

ccctagtacg tgtttagaat aaaaaccagt ttgtttttca acctggacct ccttggag       1738
```

```
<210>   36
<211>   438
<212>   PRT
<213>   Homo sapiens

<400>   36

Met Gly Arg Val Glu Asp Gly Gly Thr Thr Glu Glu Leu Glu Asp Trp
1               5                   10                  15


Asp Pro Gly Thr Ser Ala Leu Pro Ala Pro Gly Ile Lys Gln Gly Pro
            20                  25                  30


Arg Glu Gln Thr Gly Thr Gly Pro Leu Ser Gln Lys Cys Trp Glu Pro
        35                  40                  45


Glu Pro Asp Ala Pro Ser Gln Pro Gly Pro Ala Leu Trp Ser Arg Gly
    50                  55                  60


Arg Ala Arg Thr Gln Ala Leu Ala Gly Gly Ser Ser Leu Gln Gln Leu
65                  70                  75                  80
```

Asp Pro Glu Asn Thr Gly Phe Ile Gly Ala Asp Thr Phe Thr Gly Leu
                85   90   95

Val His Ser His Glu Leu Pro Leu Asp Pro Ala Lys Leu Asp Met Leu
            100   105   110

Val Ala Leu Ala Gln Ser Asn Glu Gln Gly Gln Val Cys Tyr Gln Glu
        115   120   125

Leu Val Asp Leu Ile Ser Ser Lys Arg Ser Ser Phe Lys Arg Ala
    130   135   140

Ile Ala Asn Gly Gln Arg Ala Leu Pro Arg Asp Gly Pro Leu Asp Glu
145   150   155   160

Pro Gly Leu Gly Val Tyr Lys Arg Phe Val Arg Tyr Val Ala Tyr Glu
                165   170   175

Ile Leu Pro Cys Glu Val Asp Arg Arg Trp Tyr Phe Tyr Arg His Arg
            180   185   190

Ser Cys Pro Pro Pro Val Phe Met Ala Ser Val Thr Leu Ala Gln Ile
        195   200   205

Ile Val Phe Leu Cys Tyr Gly Ala Arg Leu Asn Lys Trp Val Leu Gln
    210   215   220

Thr Tyr His Pro Glu Tyr Met Lys Ser Pro Leu Val Tyr His Pro Gly
225   230   235   240

His Arg Ala Arg Ala Trp Arg Phe Leu Thr Tyr Met Phe Met His Val
                245   250   255

Gly Leu Glu Gln Leu Gly Phe Asn Ala Leu Leu Gln Leu Met Ile Gly
            260   265   270

Val Pro Leu Glu Met Val His Gly Leu Leu Arg Ile Ser Leu Leu Tyr
        275   280   285

Leu Ala Gly Val Leu Ala Gly Ser Leu Thr Val Ser Ile Thr Asp Met
    290   295   300

Arg Ala Pro Val Val Gly Gly Ser Gly Gly Val Tyr Ala Leu Cys Ser
305   310   315   320

Ala His Leu Ala Asn Val Val Met Asn Trp Ala Gly Met Arg Cys Pro

39

                        325   330   335

Tyr Lys Leu Leu Arg Met Val Leu Ala Leu Val Cys Met Ser Ser Glu
        340   345   350

Val Gly Arg Ala Val Trp Leu Arg Phe Ser Pro Pro Leu Pro Ala Ser
        355   360   365

Gly Pro Gln Pro Ser Phe Met Ala His Leu Ala Gly Ala Val Val Gly
        370   375   380

Val Ser Met Gly Leu Thr Ile Leu Arg Ser Tyr Glu Glu Arg Leu Arg
385   390   395   400

Asp Gln Cys Gly Trp Trp Val Val Leu Leu Ala Tyr Gly Thr Phe Leu
            405   410   415

Leu Phe Ala Val Phe Trp Asn Val Phe Ala Tyr Asp Leu Leu Gly Ala
            420   425   430

His Ile Pro Pro Pro Pro
        435

<210>   37
<211>   292
<212>   PRT
<213>   Homo sapiens

<400>   37

Met Asn Leu Asn Met Gly Arg Glu Met Lys Glu Glu Leu Glu Glu Glu
1   5   10   15

Glu Lys Met Arg Glu Asp Gly Gly Gly Lys Asp Arg Ala Lys Ser Lys
        20   25   30

Lys Val His Arg Ile Val Ser Lys Trp Met Leu Pro Glu Lys Ser Arg
        35   40   45

Gly Thr Tyr Leu Glu Arg Ala Asn Cys Phe Pro Pro Val Phe Ile
    50   55   60

Ile Ser Ile Ser Leu Ala Glu Leu Ala Val Phe Ile Tyr Tyr Ala Val
65   70   75   80

Trp Lys Pro Gln Lys Gln Trp Ile Thr Leu Asp Thr Gly Ile Leu Glu
            85   90   95

Ser Pro Phe Ile Tyr Ser Pro Glu Lys Arg Glu Glu Ala Trp Arg Phe
            100   105   110

Ile Ser Tyr Met Leu Val His Ala Gly Val Gln His Ile Leu Gly Asn
            115   120   125

Leu Cys Met Gln Leu Val Leu Gly Ile Pro Leu Glu Met Val His Lys
    130   135   140

Gly Leu Arg Val Gly Leu Val Tyr Leu Ala Gly Val Ile Ala Gly Ser
145   150   155   160

Leu Ala Ser Ser Ile Phe Asp Pro Leu Arg Tyr Leu Val Gly Ala Ser
                165   170   175

Gly Gly Val Tyr Ala Leu Met Gly Gly Tyr Phe Met Asn Val Leu Val
            180   185   190

Asn Phe Gln Glu Met Ile Pro Ala Phe Gly Ile Phe Arg Leu Leu Ile
        195   200   205

Ile Ile Leu Ile Ile Val Leu Asp Met Gly Phe Ala Leu Tyr Arg Arg
    210   215   220

Phe Phe Val Pro Glu Asp Gly Ser Pro Val Ser Phe Ala Ala His Ile
225   230   235   240

Ala Gly Gly Phe Ala Gly Met Ser Ile Gly Tyr Thr Val Phe Ser Cys
            245   250   255

Phe Asp Lys Ala Leu Leu Lys Asp Pro Arg Phe Trp Ile Ala Ile Ala
        260   265   270

Ala Tyr Leu Ala Cys Val Leu Phe Ala Val Phe Phe Asn Ile Phe Leu
    275   280   285

Ser Pro Ala Asn
    290

<210> 38
<211> 174
<212> PRT
<213> Homo sapiens

<400> 38

Leu Cys Arg Val Gln His Ile Leu Gly Asn Leu Cys Met Gln Leu Val
1   5   10   15

41

Leu Gly Ile Pro Leu Glu Met Val His Lys Gly Leu Arg Val Gly Leu
              20   25   30

Val Tyr Leu Ala Gly Val Ile Ala Gly Ser Leu Ala Ser Ser Ile Phe
              35   40   45

Asp Pro Leu Arg Tyr Leu Val Gly Ala Ser Gly Gly Val Tyr Ala Leu
      50   55   60

Met Gly Gly Tyr Phe Met Asn Val Leu Val Asn Phe Gln Glu Met Ile
65   70   75   80

Pro Ala Phe Gly Ile Phe Arg Leu Leu Ile Ile Ile Leu Ile Ile Val
              85   90   95

Leu Asp Met Gly Phe Ala Leu Tyr Arg Arg Phe Phe Val Pro Glu Asp
              100   105   110

Gly Ser Pro Val Ser Phe Ala Ala His Ile Ala Gly Gly Phe Ala Gly
      115   120   125

Met Ser Ile Gly Tyr Thr Val Phe Ser Cys Phe Asp Lys Ala Leu Leu
      130   135   140

Lys Asp Pro Arg Phe Trp Ile Ala Ile Ala Ala Tyr Leu Ala Cys Val
145   150   155   160

Leu Phe Ala Val Phe Phe Asn Ile Phe Leu Ser Pro Ala Asn
              165   170

<210> 39
<211> 162
<212> PRT
<213> Homo sapiens

<400> 39

Met Gln Leu Val Leu Gly Ile Pro Leu Glu Met Val His Lys Gly Leu
1   5   10   15

Arg Val Gly Leu Val Tyr Leu Ala Gly Val Ile Ala Gly Ser Leu Ala
              20   25   30

Ser Ser Ile Phe Asp Pro Leu Arg Tyr Leu Val Gly Ala Ser Gly Gly
              35   40   45

Val Tyr Ala Leu Met Gly Gly Tyr Phe Met Asn Val Leu Val Asn Phe
      50   55   60

42

```
Gln Glu Met Ile Pro Ala Phe Gly Ile Phe Arg Leu Leu Ile Ile Ile
65              70          75                  80


Leu Ile Ile Val Leu Asp Met Gly Phe Ala Leu Tyr Arg Arg Phe Phe
                85              90              95


Val Pro Glu Asp Gly Ser Pro Val Ser Phe Ala Ala His Ile Ala Gly
            100             105             110


Gly Phe Ala Gly Met Ser Ile Gly Tyr Thr Val Phe Ser Cys Phe Asp
        115             120             125


Lys Ala Leu Leu Lys Asp Pro Arg Phe Trp Ile Ala Ile Ala Ala Tyr
    130             135             140


Leu Ala Cys Val Leu Phe Ala Val Phe Phe Asn Ile Phe Leu Ser Pro
145             150             155             160


Ala Asn



<210>   40
<211>   379
<212>   PRT
<213>   Homo sapiens

<400>   40

Met Ala Trp Arg Gly Trp Ala Gln Arg Gly Trp Gly Cys Gly Gln Ala
1               5               10              15


Trp Gly Ala Ser Val Gly Gly Arg Ser Cys Glu Glu Leu Thr Ala Val
            20              25              30


Leu Thr Pro Pro Gln Leu Leu Gly Arg Arg Phe Asn Phe Phe Ile Gln
        35              40              45


Gln Lys Cys Gly Phe Arg Lys Ala Pro Arg Lys Val Glu Pro Arg Arg
    50              55              60


Ser Asp Pro Gly Thr Ser Gly Glu Ala Tyr Lys Arg Ser Ala Leu Ile
65              70              75              80


Pro Pro Val Glu Glu Thr Val Phe Tyr Pro Ser Pro Tyr Pro Ile Arg
            85              90              95


Ser Leu Ile Lys Pro Leu Phe Phe Thr Val Gly Phe Thr Gly Cys Ala
```

                      100   105   110

Phe Gly Ser Ala Ala Ile Trp Gln Tyr Glu Ser Leu Lys Ser Arg Val
        115   120   125

Gln Ser Tyr Phe Asp Gly Ile Lys Ala Asp Trp Leu Asp Ser Ile Arg
    130   135   140

Pro Gln Lys Glu Gly Asp Phe Arg Lys Glu Ile Asn Lys Trp Trp Asn
145   150   155   160

Asn Leu Ser Asp Gly Gln Arg Thr Val Thr Gly Ile Ile Ala Ala Asn
            165   170   175

Val Leu Val Phe Cys Leu Trp Arg Val Pro Ser Leu Gln Arg Thr Met
        180   185   190

Ile Arg Tyr Phe Thr Ser Asn Pro Ala Ser Lys Val Leu Cys Ser Pro
        195   200   205

Met Leu Leu Ser Thr Phe Ser His Phe Ser Leu Phe His Met Ala Ala
    210   215   220

Asn Met Tyr Val Leu Trp Ser Phe Ser Ser Ser Ile Val Asn Ile Leu
225   230   235   240

Gly Gln Glu Gln Phe Met Ala Val Tyr Leu Ser Ala Gly Val Ile Ser
            245   250   255

Asn Phe Val Ser Tyr Leu Gly Lys Val Ala Thr Gly Arg Tyr Gly Pro
            260   265   270

Ser Leu Gly Ala Ser Gly Ala Ile Met Thr Val Leu Ala Ala Val Cys
        275   280   285

Thr Lys Ile Pro Glu Gly Arg Leu Ala Ile Ile Phe Leu Pro Met Phe
    290   295   300

Thr Phe Thr Ala Gly Asn Ala Leu Lys Ala Ile Ile Ala Met Asp Thr
305   310   315   320

Ala Gly Met Ile Leu Gly Trp Lys Phe Phe Asp His Ala Ala His Leu
            325   330   335

Gly Gly Ala Leu Phe Gly Ile Trp Tyr Val Thr Tyr Gly His Glu Leu
            340   345   350

Ile Trp Lys Asn Arg Glu Pro Leu Val Lys Ile Trp His Glu Ile Arg
        355   360   365

Thr Asn Gly Pro Lys Lys Gly Gly Gly Ser Lys
    370   375

<210> 41
<211> 855
<212> PRT
<213> Homo sapiens

<400> 41

Met Ser Glu Ala Arg Arg Asp Ser Thr Ser Ser Leu Gln Arg Lys Lys
1   5   10   15

Pro Pro Trp Leu Lys Leu Asp Ile Pro Ser Ala Val Pro Leu Thr Ala
            20   25   30

Glu Glu Pro Ser Phe Leu Gln Pro Leu Arg Arg Gln Ala Phe Leu Arg
        35   40   45

Ser Val Ser Met Pro Ala Glu Thr Ala His Ile Ser Ser Pro His His
    50   55   60

Glu Leu Arg Arg Pro Val Leu Gln Arg Gln Thr Ser Ile Thr Gln Thr
65   70   75   80

Ile Arg Arg Gly Thr Ala Asp Trp Phe Gly Val Ser Lys Asp Ser Asp
            85   90   95

Ser Thr Gln Lys Trp Gln Arg Lys Ser Ile Arg His Cys Ser Gln Arg
            100   105   110

Tyr Gly Lys Leu Lys Pro Gln Val Leu Arg Glu Leu Asp Leu Pro Ser
    115   120   125

Gln Asp Asn Val Ser Leu Thr Ser Thr Glu Thr Pro Pro Pro Leu Tyr
    130   135   140

Val Gly Pro Cys Gln Leu Gly Met Gln Lys Ile Ile Asp Pro Leu Ala
145   150   155   160

Arg Gly Arg Ala Phe Arg Val Ala Asp Asp Thr Ala Glu Gly Leu Ser
            165   170   175

Ala Pro His Thr Pro Val Thr Pro Gly Ala Ala Ser Leu Cys Ser Phe
            180   185   190

45

Ser Ser Ser Arg Ser Gly Phe His Arg Leu Pro Arg Arg Arg Lys Arg
        195   200   205

Glu Ser Val Ala Lys Met Ser Phe Arg Ala Ala Ala Ala Leu Met Lys
    210   215   220

Gly Arg Ser Val Arg Asp Gly Thr Phe Arg Arg Ala Gln Arg Arg Ser
225   230   235   240

Phe Thr Pro Ala Ser Phe Leu Glu Glu Asp Thr Thr Asp Phe Pro Asp
            245   250   255

Glu Leu Asp Thr Ser Phe Phe Ala Arg Glu Gly Ile Leu His Glu Glu
            260   265   270

Leu Ser Thr Tyr Pro Asp Glu Val Phe Glu Ser Pro Ser Glu Ala Ala
        275   280   285

Leu Lys Asp Trp Glu Lys Ala Pro Glu Gln Ala Asp Leu Thr Gly Gly
    290   295   300

Ala Leu Asp Arg Ser Glu Leu Glu Arg Ser His Leu Met Leu Pro Leu
305   310   315   320

Glu Arg Gly Trp Arg Lys Gln Lys Glu Gly Ala Ala Ala Pro Gln Pro
            325   330   335

Lys Val Arg Leu Arg Gln Glu Val Val Ser Thr Ala Gly Pro Arg Arg
        340   345   350

Gly Gln Arg Ile Ala Val Pro Val Arg Lys Leu Phe Ala Arg Glu Lys
        355   360   365

Arg Pro Tyr Gly Leu Gly Met Val Gly Arg Leu Thr Asn Arg Thr Tyr
    370   375   380

Arg Lys Arg Ile Asp Ser Phe Val Lys Arg Gln Ile Glu Asp Met Asp
385   390   395   400

Asp His Arg Pro Phe Phe Thr Tyr Trp Leu Thr Phe Val His Ser Leu
            405   410   415

Val Ala Ile Leu Ala Val Cys Ile Tyr Gly Ile Ala Pro Val Gly Phe
        420   425   430

Ser Gln His Glu Thr Val Asp Ser Val Leu Arg Asn Arg Gly Val Tyr
        435   440   445

Glu Asn Val Lys Tyr Val Gln Gln Glu Asn Phe Trp Ile Gly Pro Ser
        450   455   460

Ser Glu Ala Leu Ile His Leu Gly Ala Lys Phe Ser Pro Cys Met Arg
465   470   475   480

Gln Asp Pro Gln Val His Ser Phe Ile Arg Ser Ala Arg Glu Arg Glu
                485   490   495

Lys His Ser Ala Cys Cys Val Arg Asn Asp Arg Ser Gly Cys Val Gln
            500   505   510

Thr Ser Glu Glu Glu Cys Ser Ser Thr Leu Ala Val Trp Val Lys Trp
        515   520   525

Pro Ile His Pro Ser Ala Pro Glu Leu Ala Gly His Lys Arg Gln Phe
        530   535   540

Gly Ser Val Cys His Gln Asp Pro Arg Val Cys Asp Glu Pro Ser Ser
545   550   555   560

Glu Asp Pro His Glu Trp Pro Glu Asp Ile Thr Lys Trp Pro Ile Cys
                565   570   575

Thr Lys Asn Ser Ala Gly Asn His Thr Asn His Pro His Met Asp Cys
            580   585   590

Val Ile Thr Gly Arg Pro Cys Cys Ile Gly Thr Lys Gly Arg Cys Glu
        595   600   605

Ile Thr Ser Arg Glu Tyr Cys Asp Phe Met Arg Gly Tyr Phe His Glu
        610   615   620

Glu Ala Thr Leu Cys Ser Gln Val His Cys Met Asp Asp Val Cys Gly
625   630   635   640

Leu Leu Pro Phe Leu Asn Pro Glu Val Pro Asp Gln Phe Tyr Arg Leu
                645   650   655

Trp Leu Ser Leu Phe Leu His Ala Gly Ile Leu His Cys Leu Val Ser
            660   665   670

Ile Cys Phe Gln Met Thr Val Leu Arg Asp Leu Glu Lys Leu Ala Gly
            675   680   685

47

```
Trp His Arg Ile Ala Ile Ile Tyr Leu Leu Ser Gly Val Thr Gly Asn
    690                 695             700


Leu Ala Ser Ala Ile Phe Leu Pro Tyr Arg Ala Glu Val Gly Pro Ala
705                 710             715                     720


Gly Ser Gln Phe Gly Ile Leu Ala Cys Leu Phe Val Glu Leu Phe Gln
            725                 730             735


Ser Trp Gln Ile Leu Ala Arg Pro Trp Arg Ala Phe Phe Lys Leu Leu
            740                 745             750


Ala Val Val Leu Phe Leu Phe Thr Phe Gly Leu Leu Pro Trp Ile Asp
        755                 760                 765


Asn Phe Ala His Ile Ser Gly Phe Ile Ser Gly Leu Phe Leu Ser Phe
    770                 775                 780


Ala Phe Leu Pro Tyr Ile Ser Phe Gly Lys Phe Asp Leu Tyr Arg Lys
785                 790                 795                     800


Arg Cys Gln Ile Ile Ile Phe Gln Val Val Phe Leu Gly Leu Leu Ala
            805                 810                 815


Gly Leu Val Val Leu Phe Tyr Val Tyr Pro Val Arg Cys Glu Trp Cys
            820                 825                 830


Glu Phe Leu Thr Cys Ile Pro Phe Thr Asp Lys Phe Cys Glu Lys Tyr
            835                 840                 845


Glu Leu Asp Ala Gln Leu His
    850                 855


<210>   42
<211>   619
<212>   PRT
<213>   Homo sapiens

<400>   42

Met Ser Val Ala His Met Ser Leu Gln Ala Ala Ala Ala Leu Leu Lys
1                   5                   10                  15


Gly Arg Ser Val Leu Asp Ala Thr Gly Gln Arg Cys Arg Val Val Lys
            20                  25                  30


Arg Ser Phe Ala Phe Pro Ser Phe Leu Glu Glu Asp Val Val Asp Gly
```

```
             35                     40                      45


Ala Asp Thr Phe Asp Ser Ser Phe Phe Ser Lys Glu Glu Met Ser Ser
    50              55              60


Met Pro Asp Asp Val Phe Glu Ser Pro Pro Leu Ser Ala Ser Tyr Phe
65              70              75              80


Arg Gly Ile Pro His Ser Ala Ser Pro Val Ser Pro Asp Gly Val Gln
                85              90              95


Ile Pro Leu Lys Glu Tyr Gly Arg Ala Pro Val Pro Gly Pro Arg Arg
            100             105             110


Gly Lys Arg Ile Ala Ser Lys Val Lys His Phe Ala Phe Asp Arg Lys
        115             120             125


Lys Arg His Tyr Gly Leu Gly Val Val Gly Asn Trp Leu Asn Arg Ser
    130             135             140


Tyr Arg Arg Ser Ile Ser Ser Thr Val Gln Arg Gln Leu Glu Ser Phe
145             150             155             160


Asp Ser His Arg Pro Tyr Phe Thr Tyr Trp Leu Thr Phe Val His Val
            165             170             175


Ile Ile Thr Leu Leu Val Ile Cys Thr Tyr Gly Ile Ala Pro Val Gly
        180             185             190


Phe Ala Gln His Val Thr Thr Gln Leu Val Leu Arg Asn Lys Gly Val
        195             200             205


Tyr Glu Ser Val Lys Tyr Ile Gln Gln Glu Asn Phe Trp Val Gly Pro
    210             215             220


Ser Ser Ile Asp Leu Ile His Leu Gly Ala Lys Phe Ser Pro Cys Ile
225             230             235             240


Arg Lys Asp Gly Gln Ile Glu Gln Leu Val Leu Arg Glu Arg Asp Leu
            245             250             255


Glu Arg Asp Ser Gly Cys Cys Val Gln Asn Asp His Ser Gly Cys Ile
            260             265             270


Gln Thr Gln Arg Lys Asp Cys Ser Glu Thr Leu Ala Thr Phe Val Lys
        275             280             285
```

Trp Gln Asp Asp Thr Gly Pro Pro Met Asp Lys Ser Asp Leu Gly Gln
    290   295   300

Lys Arg Thr Ser Gly Ala Val Cys His Gln Asp Pro Arg Thr Cys Glu
305   310   315   320

Glu Pro Ala Ser Ser Gly Ala His Ile Trp Pro Asp Asp Ile Thr Lys
            325   330   335

Trp Pro Ile Cys Thr Glu Gln Ala Arg Ser Asn His Thr Gly Phe Leu
            340   345   350

His Met Asp Cys Glu Ile Lys Gly Arg Pro Cys Cys Ile Gly Thr Lys
        355   360   365

Gly Ser Cys Glu Ile Thr Thr Arg Glu Tyr Cys Glu Phe Met His Gly
    370   375   380

Tyr Phe His Glu Glu Ala Thr Leu Cys Ser Gln Val His Cys Leu Asp
385   390   395   400

Lys Val Cys Gly Leu Leu Pro Phe Leu Asn Pro Glu Val Pro Asp Gln
            405   410   415

Phe Tyr Arg Leu Trp Leu Ser Leu Phe Leu His Ala Gly Val Val His
            420   425   430

Cys Leu Val Ser Val Val Phe Gln Met Thr Ile Leu Arg Asp Leu Glu
        435   440   445

Lys Leu Ala Gly Trp His Arg Ile Ala Ile Ile Phe Ile Leu Ser Gly
    450   455   460

Ile Thr Gly Asn Leu Ala Ser Ala Ile Phe Leu Pro Tyr Arg Ala Glu
465   470   475   480

Val Gly Pro Ala Gly Ser Gln Phe Gly Leu Leu Ala Cys Leu Phe Val
            485   490   495

Glu Leu Phe Gln Ser Trp Pro Leu Leu Glu Arg Pro Trp Lys Ala Phe
            500   505   510

Leu Asn Leu Ser Ala Ile Val Leu Phe Leu Phe Ile Cys Gly Leu Leu
        515   520   525

Pro Trp Ile Asp Asn Ile Ala His Ile Phe Gly Phe Leu Ser Gly Leu

          530   535   540

Leu Leu Ala Phe Ala Phe Leu Pro Tyr Ile Thr Phe Gly Thr Ser Asp
545   550   555   560

Lys Tyr Arg Lys Arg Ala Leu Ile Leu Val Ser Leu Leu Ala Phe Ala
                565   570   575

Gly Leu Phe Ala Ala Leu Val Leu Trp Leu Tyr Ile Tyr Pro Ile Asn
                580   585   590

Trp Pro Trp Ile Glu His Leu Thr Cys Phe Pro Phe Thr Ser Arg Phe
            595   600   605

Cys Glu Lys Tyr Glu Leu Asp Gln Val Leu His
    610   615

<210>   43
<211>   434
<212>   PRT
<213>   Homo sapiens

<400>   43

Met Gly Arg Gly Leu Trp Glu Ala Trp Pro Pro Ala Gly Ser Ser Ala
1   5   10   15

Val Ala Lys Gly Asn Cys Arg Glu Glu Ala Glu Gly Ala Glu Asp Arg
            20   25   30

Gln Pro Ala Ser Arg Arg Gly Ala Gly Thr Thr Ala Ala Met Ala Ala
        35   40   45

Ser Gly Pro Gly Cys Arg Ser Trp Cys Leu Cys Pro Glu Val Pro Ser
    50   55   60

Ala Thr Phe Phe Thr Ala Leu Leu Ser Leu Leu Val Ser Gly Pro Arg
65   70   75   80

Leu Phe Leu Leu Gln Gln Pro Leu Ala Pro Ser Gly Leu Thr Leu Lys
                85   90   95

Ser Glu Ala Leu Arg Asn Trp Gln Val Tyr Arg Leu Val Thr Tyr Ile
            100   105   110

Phe Val Tyr Glu Asn Pro Ile Ser Leu Leu Cys Gly Ala Ile Ile Ile
            115   120   125

51

Trp Arg Phe Ala Gly Asn Phe Glu Arg Thr Val Gly Thr Val Arg His
    130   135   140

Cys Phe Phe Thr Val Ile Phe Ala Ile Phe Ser Ala Ile Ile Phe Leu
145   150   155   160

Ser Phe Glu Ala Val Ser Ser Leu Ser Lys Leu Gly Glu Val Glu Asp
                165   170   175

Ala Arg Gly Phe Thr Pro Val Ala Phe Ala Met Leu Gly Val Thr Thr
            180   185   190

Val Arg Ser Arg Met Arg Arg Ala Leu Val Phe Gly Met Val Val Pro
            195   200   205

Ser Val Leu Val Pro Trp Leu Leu Leu Gly Ala Ser Trp Leu Ile Pro
    210   215   220

Gln Thr Ser Phe Leu Ser Asn Val Cys Gly Leu Ser Ile Gly Leu Ala
225   230   235   240

Tyr Ala His Leu Leu Leu Phe His Arg Pro Leu Arg Ala Ser Gly Ala
                245   250   255

Glu Ala Arg Ser Asp Leu Pro Leu Gln Pro Asp Glu Glu Asp Ile Arg
            260   265   270

Val Gln Val Arg Leu Arg Val Phe Ser Arg Glu Glu Gly Ser Pro Glu
            275   280   285

Pro Glu Thr Glu Pro Gly Ala Trp Leu Leu Pro His Thr Glu Leu Pro
    290   295   300

Pro Ser Pro Val Pro Lys Pro Pro Cys Val Pro Asp Ala Ala Arg Gln
305   310   315   320

Trp Ser Glu Ala Gly Leu Leu Ala Ser Cys Thr Pro Gly His Met Pro
            325   330   335

Thr Leu Pro Pro Tyr Gln Pro Ala Ser Gly Leu Cys Tyr Val Gln Asn
            340   345   350

His Phe Gly Pro Asn Pro Thr Ser Ser Ser Val Tyr Pro Ala Ser Ala
        355   360   365

Gly Thr Ser Leu Gly Ile Gln Pro Pro Thr Pro Val Asn Ser Pro Gly
    370   375   380

52

Thr Val Tyr Ser Gly Ala Leu Gly His Gln Gly Leu Gln Ala Pro Arg
385   390   395   400

Ser Pro Pro Gly Ser Pro Cys Pro Glu Arg Ile Ser Arg Glu Val Ile
                405   410   415

Ser Leu Gly Leu Leu Lys Val Leu Pro Lys Ser Leu Leu Thr Lys Val
            420   425   430

Thr Tyr

<210> 44
<211> 386
<212> PRT
<213> Homo sapiens

<400> 44

Met His Ala Arg Gly Pro His Gly Gln Leu Ser Pro Ala Leu Pro Leu
1   5   10   15

Ala Ser Ser Val Leu Met Leu Leu Met Ser Thr Leu Trp Leu Val Gly
            20   25   30

Ala Gly Pro Gly Leu Val Leu Ala Pro Glu Leu Leu Leu Asp Pro Trp
        35   40   45

Gln Val His Arg Leu Leu Thr His Ala Leu Gly His Thr Ala Leu Pro
    50   55   60

Gly Leu Leu Leu Ser Leu Leu Leu Pro Thr Val Gly Trp Gln Gln
65   70   75   80

Glu Cys His Leu Gly Thr Leu Arg Phe Leu His Ala Ser Ala Leu Leu
                85   90   95

Ala Leu Ala Ser Gly Leu Leu Ala Val Leu Leu Ala Gly Leu Gly Leu
            100   105   110

Ser Ser Ala Ala Gly Ser Cys Gly Tyr Met Pro Val His Leu Ala Met
            115   120   125

Leu Ala Gly Glu Gly His Arg Pro Arg Arg Pro Arg Gly Ala Leu Pro
        130   135   140

Pro Trp Leu Ser Pro Trp Leu Leu Leu Ala Leu Thr Pro Leu Leu Ser

```
145                  150                 155                 160
```

Ser Glu Pro Pro Phe Leu Gln Leu Leu Cys Gly Leu Leu Ala Gly Leu
            165   170   175

Ala Tyr Ala Ala Gly Ala Phe Arg Trp Leu Glu Pro Ser Glu Arg Arg
            180   185   190

Leu Gln Val Leu Gln Glu Gly Val Leu Cys Arg Thr Leu Ala Gly Cys
            195   200   205

Trp Pro Leu Arg Leu Leu Ala Thr Pro Gly Ser Leu Ala Glu Leu Pro
    210   215   220

Val Thr His Pro Ala Gly Val Arg Pro Pro Ile Pro Gly Pro Pro Tyr
225   230   235   240

Val Ala Ser Pro Asp Leu Trp Ser His Trp Glu Asp Ser Ala Leu Pro
            245   250   255

Pro Pro Ser Leu Arg Pro Val Gln Pro Thr Trp Glu Gly Ser Ser Glu
            260   265   270

Ala Gly Leu Asp Trp Ala Gly Ala Ser Phe Ser Pro Gly Thr Pro Met
            275   280   285

Trp Ala Ala Leu Asp Glu Gln Met Leu Gln Glu Gly Ile Gln Ala Ser
    290   295   300

Leu Leu Asp Gly Pro Ala Gln Glu Pro Gln Ser Ala Pro Trp Leu Ser
305   310   315   320

Lys Ser Ser Val Ser Ser Leu Arg Leu Gln Gln Leu Glu Arg Met Gly
            325   330   335

Phe Pro Thr Glu Gln Ala Val Val Ala Leu Ala Ala Thr Gly Arg Val
            340   345   350

Glu Gly Ala Val Ser Leu Leu Val Gly Gly Gln Val Gly Thr Glu Thr
            355   360   365

Leu Val Thr His Gly Lys Gly Gly Pro Ala His Ser Glu Gly Pro Gly
    370   375   380

Pro Pro
385

54

```
<210>   45
<211>   1413
<212>   DNA
<213>   Mus musculus

<400>   45
atggcctact gggagaagct tctgccgcac gtggtgcaag agctggacgc aggcagtggt      60

tctgggtcga gctgttcggt gggggacgat ctacccaggc agtgcgtcga gttggggccg     120

cgcgtgcgcg tgcaggatat gcgcgtactc tggctgtgtg cgcgctgggc tggcatctcg     180

gtgacgctgt ctgagctcgg ggagcgcaga gcagagcagg cggagcgggc cggccggggc     240

ggagcggagc ggtccgcaga gcagcccctc ccggcctcgg ccgaccccgg ccctcggccc     300

ggctctatgg acaggagctc gctgctgcag ctcattcagg agcagcagct ggatcctgag     360

aacacaggct tcatcggtgc ggacaccttc gctggtctgg tacacagcca tgagctgccc     420

ctggacccca ccaagttgga catgttggtg gctctggctc agagcaacga gcggggccag     480

gtctgctacc aggagctggt ggacctggtc agtgccatga tcagcagcaa gcgttccagc     540

agcttcaaga gagccattgc taacggacag cgggcactgc cccgagacgg actgctagat     600

gagccaggcc tgagtgtcta caagcggttt gtgcgctatg tggcctacga gatcctgccc     660

tgcgaggtgg accgccgctg gtacttctac aggcaccgca cctgcccacc ccctgtgttc     720

atggcctcgg tcactcttgc ccagtgtccc cagatcatcg tgttcctgtg ctacggggca     780

cgtctcaaca gtgggtgct ccagacctac cacctgaat atatgaagag ccctctggtg     840

tatcacccag dacaccgtgc tcgggcatgg cgcttcctca cctacatgtt catgcatgtc     900

gggctggagc agctagggtt caatgccctc ctgcagctga tgatcggtgt gcccctggag     960

atggtacatg gcgtgcttcg catcagcctg ctctacctgg cgggcgtgct ggcaggctcc    1020

ctgactgtct ctatcacaga tatgcgtgcc cccgtggtag ggggctctgg aggggtctat.    1080

gccctgtgct cagcacacct ggccaatgtt gtcatggtaa tgggactgcc tttctgggga    1140

gtgagttccg aagtgggccg ggctgtgtgg ctacgcttct ccccaccact gcctgcctca    1200

ggcccacagc ccagcttcat ggcacacctg gctggtgcag tggtaggtgt aagcatgggc    1260

cttaccatcc ttcggagcta tgaggaacgc ctgagggacc agtgcggctg gtgggtggtg    1320

ctacttgcct atggcacctt cctgcttttc gccatcttct ggaacgtctt tgcctatgac    1380

ctgctgggtg ccgatatccc ccctccacct tga                                 1413

<210>   46
<211>   470
<212>   PRT
<213>   Mus musculus

<400>   46
```

Met Ala Tyr Trp Glu Lys Leu Leu Pro His Val Val Gln Glu Leu Asp
1   5   10   15

Ala Gly Ser Gly Ser Gly Ser Ser Cys Ser Val Gly Asp Asp Leu Pro
            20   25   30

Arg Gln Cys Val Glu Leu Gly Pro Arg Val Arg Val Gln Asp Met Arg
        35   40   45

Val Leu Trp Leu Cys Ala Arg Trp Ala Gly Ile Ser Val Thr Leu Ser
    50   55   60

Glu Leu Gly Glu Arg Arg Ala Glu Gln Ala Glu Arg Ala Gly Arg Gly
65   70   75   80

Gly Ala Glu Arg Ser Ala Glu Gln Pro Leu Pro Ala Ser Ala Asp Pro
            85   90   95

Gly Pro Arg Pro Gly Ser Met Asp Arg Ser Ser Leu Leu Gln Leu Ile
            100   105   110

Gln Glu Gln Gln Leu Asp Pro Glu Asn Thr Gly Phe Ile Gly Ala Asp
        115   120   125

Thr Phe Ala Gly Leu Val His Ser His Glu Leu Pro Leu Asp Pro Thr
    130   135   140

Lys Leu Asp Met Leu Val Ala Leu Ala Gln Ser Asn Glu Arg Gly Gln
145   150   155   160

Val Cys Tyr Gln Glu Leu Val Asp Leu Val Ser Ala Met Ile Ser Ser
            165   170   175

Lys Arg Ser Ser Ser Phe Lys Arg Ala Ile Ala Asn Gly Gln Arg Ala
            180   185   190

Leu Pro Arg Asp Gly Leu Leu Asp Glu Pro Gly Leu Ser Val Tyr Lys
        195   200   205

Arg Phe Val Arg Tyr Val Ala Tyr Glu Ile Leu Pro Cys Glu Val Asp
    210   215   220

Arg Arg Trp Tyr Phe Tyr Arg His Arg Thr Cys Pro Pro Pro Val Phe
225   230   235   240

Met Ala Ser Val Thr Leu Ala Gln Cys Pro Gln Ile Ile Val Phe Leu

56

                    245   250   255

    Cys Tyr Gly Ala Arg Leu Asn Lys Trp Val Leu Gln Thr Tyr His Pro
            260   265   270

    Glu Tyr Met Lys Ser Pro Leu Val Tyr His Pro Gly His Arg Ala Arg
            275   280   285

    Ala Trp Arg Phe Leu Thr Tyr Met Phe Met His Val Gly Leu Glu Gln
        290   295   300

    Leu Gly Phe Asn Ala Leu Leu Gln Leu Met Ile Gly Val Pro Leu Glu
    305   310   315   320

    Met Val His Gly Val Leu Arg Ile Ser Leu Leu Tyr Leu Ala Gly Val
                325   330   335

    Leu Ala Gly Ser Leu Thr Val Ser Ile Thr Asp Met Arg Ala Pro Val
            340   345   350

    Val Gly Gly Ser Gly Gly Val Tyr Ala Leu Cys Ser Ala His Leu Ala
            355   360   365

    Asn Val Val Met Val Met Gly Leu Pro Phe Trp Gly Val Ser Ser Glu
        370   375   380

    Val Gly Arg Ala Val Trp Leu Arg Phe Ser Pro Pro Leu Pro Ala Ser
    385   390   395   400

    Gly Pro Gln Pro Ser Phe Met Ala His Leu Ala Gly Ala Val Val Gly
                405   410   415

    Val Ser Met Gly Leu Thr Ile Leu Arg Ser Tyr Glu Glu Arg Leu Arg
                420   425   430

    Asp Gln Cys Gly Trp Trp Val Val Leu Leu Ala Tyr Gly Thr Phe Leu
            435   440   445

    Leu Phe Ala Ile Phe Trp Asn Val Phe Ala Tyr Asp Leu Leu Gly Ala
        450   455   460

    Asp Ile Pro Pro Pro Pro
    465   470

    <210> 47
    <211> 34319
    <212> DNA

<213> Mus musculus

<220>
<221> misc_feature
<222> (34274)..(34275)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (34285)..(34285)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (34297)..(34298)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (34301)..(34301)
<223> n is a, c, g, or t

<400> 47
actagtcctg caggtttaaa cgaattcgcc ctttccctcc tgagcctgcc ctcacccttg   60

tccagtgttc agccaagcac ttgaggaata tctcagagct gttctactat gcccagaagg   120

ctgttctgca tcccacagcc cccctttatg accctgaggc caaacaggta agtagtggat   180

gagggccaca cacctatttc tcgaaggccc atccatgtgg gccttggata aggagtacaa   240

gtaaatacgc gtaccgacag ggtggccagg aaaagggcca cttggtgttg tgtccctgca   300

gctcaggcct gcatgcgccc aggcgctcac acgcatcttc aggctgtcgg accaggaccg   360

agaccacggg ctgagcgacg aggaactcaa tgccttccag gtgtggtcct gccctgacac   420

cctagctgcc cttgtgagag ggtcatcagg atagctgact gatctgaagc actctcttac   480

cggaggcaga agtcctgctt tggccatccc ctggccccac aggccctgga ggatgtgaag   540

agggtggtgt gcaagaacgt gtcaggtggt gtgcagaatg accggctgac cctggaaggt   600

gaggcagaca tcaccttata tcctgatgtg tggggagtgg caaggcggtg tactaacctt   660

ccaccttccc acaggcttcc tcttcctgaa cacactcttc atccagcgtg gccggcacga   720

gaccacgtgg accatcctgc ggcgctttgg ttacagtgac tcactggagc taacgcccga   780

ctacctctat ccagcgtgag ttgtaggctg gtgggcagtg gtgcttatcc tcttggggtc   840

aatccttgtg gcctggggac tttagcttgg tctctgagca cacactcatt agctcccccc   900

ttgcttagcc ccaaccgtgt aaccagaggt tctgttggga gtgcctgact gcccaggcac   960

ttttaactgt cacaggtgac agtaaccctg ccttgtgaca gttaaagtgc atccccacag 1020

gctccatgtc cccctggct gcagcacaga gctcaaccat cgtggctacc agtttgtgca 1080

gcggatgttc gagaagcatg accaggtaag cacgtcacct cacgccttga cacccctccc 1140

```
cgcaaagaaa cttgtcacca tagccccttt gcctgtagga ccatgatggt gtcctctcgc   1200

ccacggaact acagaatctc ttcagtgtgt tctcaggtgc tccatggggc cctgaactct   1260

tacacacagt ccccactcag gctggctgtc tgcctttgca tggctatctc tgccagtgga   1320

cgtaagtaca gtccctgtct acctgtctgt ccctacccat cccatcttcc actctcctta   1380

gccctgtacc ctcctctact cccagcctca tgacctactt ggatgtccag caatgccttg   1440

cacaccttgg ctacctgggt taccctaccc tctgtgaaca ggactctcag gcacaggcta   1500

tcacaggtag gtgaccactt ttttggtccc tgggcccagc caaacccttc ttgatggcca   1560

acaccctcac gtgagcgcac cttttcagtt acccgcgaga agaagctgga ccaggagaaa   1620

gggcagacac agcgcagcgt tctcatgtgt aaggtgctgg gagcccgagg agtgggcaag   1680

tcagccttct tgcaagcctt ccttggcaac agccttgggg tgagaattat ggaggcccca   1740

gcccctacct gggaactcaa gggggcagag gggtgccaca gagaactctg gggccaaaaa   1800

ctcttcctgc tgtccacaca ggaagccagg gacccccctg agaaattccc cctgcatacc   1860

atcaatacag tgcgggtcaa cgggcaggag aagtacctga ttgtgagtga gggcctgtgc   1920

accatgtcct gaagggtata cctgtgacac ccatataccc atatattctt cctcctaatg   1980

gagctgtctc agggtggccc ttgagggatc cagctgtggg gaggccacta aagctactgt   2040

ttgtcccagc tgtgtgaagt gaatgctgac agcctgctgg acacctctct agacactacc   2100

tgtgatgtcg cctgcttaat gtttgacagc agtgatccca agacctttgt acactgtgct   2160

accatataca aggcaagccc tgacttgagg tcctgtgggg ggatcccagc catagcccag   2220

ttagccagca agctcagctc tcgaatgttt cttttttcct ggcagcgtta ttacatggat   2280

gggcagaccc cctgcctctt tattgcctcc aaagctgatc tccctgaagg tgttgccccc   2340

ccagggctgt cacccgctga gttctgccgg aggcaccggc tgcctgcccc agcttcattc   2400

tcctgcttgg gacccgcaat gcccagcaca gatgtcttta cccaacttgc taccatggcc   2460

actttcccgt gagtataccg cagcctcaat gtgggtgggg cttagtctgt ctgggtacaa   2520

aggtgtcaga gctgtagacc cgtggcctgt gcccagaga gctatgctgg atggcctctg   2580

gcttggcctc cgcctgcctg ggacgggttt tcagcagtga gcttttcgca agcacacgtc   2640

ttgagtatta ttctttccac agacacctgg tgcacacaga gctgcacccc acctccttct   2700

ggctccgagg agtgctggtg gctgtcggca cggctgtggc tgctgtcctt agcttctcac   2760

tatacagggt cctagtgaag agccgatgac aaaggaccca tgagcctctc ctgcccccag   2820

gcacagcaca ctgtggtgcc ttatgtgggc ccacacggca aagaagtatg ggctttcggc   2880

ctgttggagt aatatctgtg tggctggcta cagggtccct gggactgatg tgggttcctg   2940

ctcaggaact tggtatttag atcttggctt ccaaaagtct gttgccctat tccaggcctg   3000
```

59

```
tcctgctgag cctgagagta ctggtgccag gtgtgaggcc ttagccttgg agggagggag   3060

gattgtgtgt aaggggctgg gctgtgggta aggaatccag tgaccaggac ccgaagttct   3120

cagggtgtgt tccccccct tcctggttgt gcatggccag tggacatgta gagggcaaga   3180

ggtagtggtc aaatcaggga gggacccgat agtcacagag gtcccacctt tttgtcaggt   3240

gtggagcctt atccacactg cgccagctgc agagcaagcg gaactgactt tcctgattaa   3300

atttcacttg tcttttccat ctcttggatt ttgatctctc aatgcaactt actttctctc   3360

tttctctccc tcccttcctt ccttcctttc ttccttcttc ctttcttctt tctttctttc   3420

tttctttctt tctttctttc tttctttctt tcaaaaatat gcttattttt atgagtttga   3480

gtgttttgcc tgcccatata tgtaagtgcc ccgtgtgtgt gcctggtacc gacagaagcc   3540

aaatgagtat gtcagagttc cctggaactg gagttaacta atgattgtga gccatcccgt   3600

gggtgttgga aacaaaccca ggtcctctgc agaggcagca agtgctctga accaccaagt   3660

cctctctcca gccccataaa aaatggcttt tcaatcgcaa tgtgactgct gtttttgtgt   3720

gagttggaaa tgccacctga cccttctgag acgtctcccc tgtgtcctgg ttcccattct   3780

gtggtctcac tagtcatgct gattaatggc tacggaaggt ccagttaagg ggtctgctgt   3840

ctagtaatgc cagcatccat cagtgtcatg ggggtggggg tgtgtgtgag agggctcctg   3900

ctgcttctct gggaacattg cttctgctgt tggttcttca ttgccccaca ctccacacag   3960

cccaggcagc ctgcctggct ctgctgagaa cctgggcttt atgtgtagcc tggctgttac   4020

aagcaagggc cagccataga tctcaacctg gccagaagct gagtgaatga tggactcgga   4080

gttgatcctg tgacatgacc tcctaccacc aaagacaccc aaacaggaat ccatacactc   4140

aacgagtcac ttgggttggg tcaggtgggt tctctctgcc ccagaatgtc cacagaacag   4200

ggatccctga aaaatgaggc ctttgcctcc tgctccagtc ccaacctact ggcctgccgt   4260

cccactgagc atctaggcag gctgcttatt gtaggcagct gtattccctg tagaggccca   4320

tcgggcaaga ataactgcca tcccatcttt gcagtagcag attgttttga atgatacttg   4380

gggggggttgc catggttaca catcctattg gtctgtcacc ttcactgtcc tcaggagcac   4440

accccctccc ttatccaagc tgcagtcagt tctaccttca cgtggctatt tgtggctctt   4500

cctggctatc ccctgtgcct tcttgactgg tgttttgca ctggaggccc tgctatctca   4560

aatccaattc agggatgttt ggttagagct tccagccttg gccatgaata gggaacctgg   4620

ggcttcgttg ctggaacagt gcaaataagg acaagtgtgt ctactgaaag cagctccagc   4680

tcaggagaaa gtatgatggg agcaactgga ccttgacatt ctggggatgg gtcacagacc   4740

tattttttagt agaaagaaag tatcctaggg aaggtaataa aacaggctgc ctaagagggt   4800

attgggtgct gcaggctgcc agctatggcc tactgggaga agcttctgcc gcacgtggtg   4860
```

60

```
caagagctgg acgcaggcag tggttctggg tcgagctgtt cggtgggggga cgatctaccc   4920

aggcagtgcg tcgagttggg gccgcgcgtg cgcgtgcagg atatgcgcgt actctggctg   4980

tgtgcgcgct gggctggcat ctcggtgacg ctgtctgagc tcggggagcg cagagcagag   5040

caggcggagc gggccggccg gggcggagcg gagcggtccg cagagcagcc cctcccggcc   5100

tcggccgacc ccggccctcg gcccggctct atggacagga gctcgctgct gcagctcatt   5160

caggagcagg tgcgtggtgg gtgggagctg gcactgcctt ggcgcccttc cagccggtga   5220

ggggttcggg ggagggacca tgctcgcaga aggcccaggc agccgcctcc caaacccccc   5280

tccctgcacg accttggtct ctgagactcc cgaccccaag aaaggtctgg cttcatcggg   5340

tttccacgga aagctgggga tggagggtgg tggtccttgg acctaggtcc ttccccttttc   5400

taggactggt agagcgaata tgaggtggta gaacagccaa accaggcacc ttgccctgcc   5460

agggcctggg aacaagctgg ttctagcaaa caggcactga ctggcacaaa ggtccagcct   5520

tcctagccag cctagctcag ccctgtggct gggagttggc ccaccctcta ctcccccacc   5580

acagcagctg gatcctgaga acacaggctt catcggtgcg gacaccttcg ctggtctggt   5640

acacagccat gagctgcccc tggaccccac caagttggac atgttggtgg ctctggctca   5700

gagcaacgag cggggccagg tctgctacca ggagctggtg gacctggtca gtgccatggt   5760

gtgtgggcag caggggaggg cccaggcctg gtcaagcgca gcaccctgac caaagctggg   5820

tgggcacaca gcttttttact gggtagggct tacagacctg ggcagtgtgg aagaccctca   5880

gaccattagc agccacacag ccaagtaagg tagggttcca caggaacaat actgtcacct   5940

acactaaggt tcctgccctt cagatcagca gcaagcgttc cagcagcttc aagagagcca   6000

ttgctaacgg acagcgggca ctgccccgag acggactgct agatgagcca ggcctgagtg   6060

tctacaagcg gtttgtgcgc tatgtggcct acgagatcct gccctgcgag gtggaccgcc   6120

gctggtactt ctacaggcac cgcacctgcc cacccctgt gttcatggcc tcggtcactc   6180

ttgcccaggt gggcctgcct acacattgcc ctggggcctc ctgagtccct agtcttgacc   6240

agccctaaca ctagtgtccc cagatcatcg tgttcctgtg ctacggggca cgtctcaaca   6300

agtgggtgct ccagacctac caccctgaat atatgaagag ccctctggtg tatcacccag   6360

gacaccgtgc tcgggcatgg cgcttcctca cctacatgtt catgcatgtc gggtaagtga   6420

gccttgcctt ccagagcctc ttctccactc ctaccccag cacctgccca ctgactctga   6480

aggcttatct cccagctaag ttacactgcc tgaggtggtc tggagtaggc agggtaccca   6540

tctgaccacc tctgccatgc tcttctgtgg ccaggctgga gcagctaggg ttcaatgccc   6600

tcctgcagct gatgatcggt gtgcccctgg agatggtaca tggcgtgctt cgcatcagcc   6660

tgctctacct ggcgggcgtg ctggcaggtg agactggcat ataggtgggt gggtcccctg   6720
```

```
cctacctcac ctgctgaggg gagggccagc tgtgccctca cctctccctt tgcccacagg   6780

ctccctgact gtctctatca cagatatgcg tgcccccgtg gtaggggggct ctggagggggt   6840

ctatgccctg tgctcagcac acctggccaa tgttgtcatg gtaatgggac tgcctttctg   6900

gggaggtggg aagggggccca ttggacatgc ttcacagcct gtctgcctgc ccccatcaga   6960

actgggctgg gatgcggtgt ccatacaagc tgctgaggat ggtgctggct ctggtgtgca   7020

gtgagtagtg gggcaggtgg ggcgctggga ggcccttttac ctgcagacag ggcccctccc   7080

acctgttgct ccctctgcag tgagttccga agtgggccgg gctgtgtggc tacgcttctc   7140

cccaccactg cctgcctcag gcccacagcc cagcttcatg gcacacctgg ctggtgcagt   7200

ggtaggtgta agcatgggcc ttaccatcct tcggagctat gaggaacgcc tgagggacca   7260

gtgcggctgg tgggtggtgc tacttgccta tggcaccttc ctgcttttcg ccatcttctg   7320

gaacgtcttt gcctatgacc tgctgggtgc cgatatcccc cctccacctt gacctgctcc   7380

ctagggccat gtggctagag ccaggcacca ggtgggcctg catgtctgcc ctatatgaat   7440

ggacctctcg gctgcttttc cccacagggg caggggggcc ctgtgatccc ccaccccccag   7500

ctgaagccaa gtcttactgc agcccttgct gcacccccc ccccaggct ttccaggggg   7560

aggggtgtta ctgggccagg ctgggtgga ggaccttgag cgtggcccta gaggagaccc   7620

cttccctccc cactgacccc aggacttgca gtctgagcct ttttggagaa tgaataaata   7680

ttttacacag caccaggtag ctgtcctggg gctctgcagg cttgtctgcc ccctcaccca   7740

agctgtatga gacacctccc ctccaggacg gtctgtgctc cctctctccc ttcaagctct   7800

gcacctgact tgaggtggta gctttactcc tggccatggc cggaatccac aggaggtggg   7860

gcccctggga agaaactgtg gacttccagg tcagtccaag ttcagcaatc gtttattgtg   7920

ctacaagagc ttccaatctc ctgttctttc agcctcccctt ccaaacgcaa gcccacaaca   7980

gagctaggtt cacttggcct gggtcctgta attcttggca ggtccttgtg ggaggcccag   8040

tcagggtgaa ggtcaggcag caggtctggg catacatctg ccccccaagt ttgcacaaga   8100

cgaagagtct acaactctaa gacactcact agtgaggtcc acaagggttc tggggtcaag   8160

ttactaacca ccagaaccat gatgggggccc tggatgacct ctcaggcccc gctcagcccc   8220

gctcagaaag gcagtgaact cttgcacctg ggctagaaac agcaagtagt gcttgagtgt   8280

gtactggagg gttagcaagc taggaactgt agccaccgcc tcttcgtata tgaaagctgc   8340

aacccccagaa acctctgcct ccaggaccac tgcaaagagt gacaaggagg ttctggggat   8400

agcttttttt gagctcctag agcccccacc caacttgggc aaagatgcta cctactcact   8460

cgaggcagtg gcagaaacca cgcagcgaag tttgactgcg gcgcccagat cagcagaaag   8520

ccgagttctg cttctccaag tccgggggttg ggggtgttag cggccaacag ctgatgcagc   8580
```

62

```
agcggctaga aatgcctcaa cgacagacca tggggtcaaa gggtcttggt gcaggcgctc    8640

tccaaggaaa ccagccggct ccgcccttga agaagctcct ccccatgacc gcgagaaggg    8700

caagagtcca gccgtgcacc cgcagctact tctatagcaa ccgcagaagg tccttcggtg    8760

tcagcgtgtc acgtgcttgt ttctcagcca ttaatggccg caggcagctc ctcccgattc    8820

ccgaggaagc ccctcgagag cggtctgatc tgcgggcctc tcttctaccc tcaattccgc    8880

cttcctaccg gcggtccaga cgttcccagg ctgtcgcgtg agcccgtgtg tgccctggtg    8940

gtgcgctagg acccgggcct agtgaccggc cctcggaagg ccacttctca atgaccgctg    9000

ccacgttcgc ccagcttggg gccggagggc ggggcttgag cgccgatggg cggagccttg    9060

gtctgagcag ggtaggtctc ggttactgta gggcggggcc tgtagagcgc ggcggaagct    9120

tctcagagct aagggaactg gctcaatcca cgaggctccg cctttgctca cgtgtcccca    9180

tgccagctcc gccccacacc ggaagttccg gtggcggatc gccgaccggg cggagctgat    9240

cgctgcgcgg gctgcgagat ctaggtggcc gggcgcggag cccaagccgt gccgcgcggc    9300

gccatgaagg gcaaggagga aaaggagggc ggcgcgcggc tgggcactgg tggcggcggc    9360

agccctgata agagcccgag tgcgcaagag ctcaaggagc agggaaaccg gctcttcgtg    9420

ggccgcaagt acccggaggc ggcggcctgc tacggccgcg ccatcgtgag tgcgggaggg    9480

gcggggggcgg ggcccgatgt gcctcccagt tttgtgatcc cggacagggt ttctggcagc    9540

ctggactcga tggctcgctc gctcccctgg tgcccggatt gctgttcccg agtctagggc    9600

tgatttgact gggaactggc gggcgggatg cttgattaag taagactaag gatgacacca    9660

ggagctcttg ccccttgaga tttccaagcc cagagcttgt agaggggttg tgggatttct    9720

ctcaaaagca gttggaggca agaaagcatg gccagaatcc agagctcttg ataattcgga    9780

aatctccagt tcctctattc cagccatggg cccccaaagc tggagcattg accagatcat    9840

tggctttcag cttgctaaat gagaggggat caactgagcc cacactcttt gacgcctggt    9900

cttggtcccc agactcggaa cccacttgtg gcagtgtact acactaaccg ggccctgtgc    9960

tatctgaaga tgcagcagcc tgaacaggca cttgctgact gccggcgagc cctggagctg    10020

gacgggcagt ctgtgaaggc gcacttcttc ctggggcagt gccagctgga gatggagagt    10080

tatgatgagg ccattgccaa tctgcagcga ggtggctga caagatgcaa ggttgagggc    10140

accttggggg accaggattg ttgacaagtg taaccaaaag tgtccactcc cctagcctat    10200

agtttggcca aggagcagcg actcaacttt gggatgata ttcctagtgc ccttcgcatt    10260

gctaagaaga agcgctggaa cagtatcgag aacggcgca tccaccagga gagtgagctg    10320

cattcatatc tcaccaggct cattgctgct gagcgagaga ggtgggtccc tcccccagca    10380

gacactgagt caccagcaca agtgttttta ttgaaggaca acaaccggta catgggaagc    10440
```

```
ccaggaggaa agaggcctga gatggggcgg gctctgtgat ttcctggtca gggccatcac   10500

caactagctt ctggttaacc ctcagggaac tggaggagtg tcagcggaac cacgagggtc   10560

atgaagatga tggccacatc cgggcccagc aggcctgcat tgaggccaag cacgtgaggg   10620

tgcccctagt ccacatgtgg gtcattgtgt gtgtgtttgt gtgtggcatg aaaagtgccc   10680

cagcctttgg tggatctgtt ggcccatgca tgacatgagt gccctggtgg acagagtgtg   10740

tgtgtctgtg tggagacagg agttgcccac acatggcact caactctgca ggataaatac   10800

atggcagata tggatgagct cttctctcag gtggacgaga aagaaaggt gaggacccc   10860

ggaagctcat cggctgctgc tggtatggtg accccacag ccccctccag ttctagacat   10920

tttcccttg tgctatagaa gcgagatatc cctgactact tgtgtggcaa gattagcttt   10980

gagctgatgc gggaaccctg cattcaccc agtggtatca cctatgaccg caaggacatt   11040

gaggagcacc tgcaggtaag gagggcccat ggagaccggg ccatggagag ccttggtacc   11100

cggatcggta tactcaggcc attgaacact cttgactctg ttgcagcgtg tgggccactt   11160

tgaccctgtg acccggagcc ctctgaccca ggaacagctc atccccaact tggccatgaa   11220

ggaagtcatt gacgctttca tctctgagaa cggctgggta gaggactatt gaggccccat   11280

gtcctgcctg gcaccctggc ccaggaggat ctggagacgg aagctccagt ccctgtatag   11340

tttgtgtccc tgggcctgcc cccatcggcc ctgctgatgg gttctgaact gctcccttc   11400

tcagcatacc ccttgctgga ccatgagcct cccttgtccc ccttctgggc tggagagtgg   11460

gtgagggtgg gctgaggttg ctgctgctgc cactgtcctg taataaagtc tgtgagcact   11520

acattggcat gtgctggtgc agtgggcttg ccagttgcct gtctggctag ccaaggaagg   11580

tagaaatgaa gacactggtg tccagattga gtgtggcatg ccaccaccga tcaggaaaat   11640

acagtacctg ggagaaaaga aaaaaaaat tttgattcat ggcttttga gttagcccaa   11700

ccccccaccta ccagctactg acttcaccag cctggatggt acattctaga ggccgtgctg   11760

acagggctag agatgggtat atttccagca gccaggccaa tgtggtcttg ttagggtgga   11820

actcaggtgt cttctcagga gggtagagga accagcgctg gaggcaggag tgtaaggtca   11880

ctatgtgtgc aatggactat agacaccaga tttacacaca cacataacac acacataaca   11940

ctgtactgac cttccgacca tagataacct ctgagaaacc aggcccatgc cagtggaagg   12000

gtaccccaga tccagctcct gtggggtcag ttatgaatag ctgatgaaca aaatatatat   12060

gggctccccc aagtttgttt tgaagtagcc cctacctgca attccaaagc tgtaagcagg   12120

ggtggttccc aggagacgga atggtggcgg agagtagtgc tggaagaggg atgcccactc   12180

agtgaagttg ttgtctccaa aaaagtacag ggtgtctgcc agcagagagg gtgataggca   12240

ggagtactgg caacagtggg ctcccaagta ccccactccc accgctgcct gggtctggct   12300
```

64

```
caccattgcc tagggatgca ggatcctggg gctgcagcag ctgttccaca tattcctgga   12360

agggcaggtc cactacaaga aggggggagcc agtcgggact ggtcagttct gcctttgaga   12420

tgtgaaggca aaccttgcag ggtgtgtagc tcacctttct ggtaggagta ggtgttggct   12480

gtactcaagc gaacaatgtt gtccccgaac gaggctagca ggttttcccg ggaacacagg   12540

gcccggaact tctgggggatg aaggataagg taactgagca cttggccaaa cagaagcagt   12600

tgcccacagg gtacgaatac tcaccgagtt gtccgtgagt ccttgcaaga tgacgggctt   12660

gaggaaggca tagctggggg gaaagaggag aatgttgcac tgagtttgaa ctctgcaggg   12720

gcctctccta cacaccaggc agtgagacct ccgcagccaa agaccatgat tggtagtatg   12780

cctagaatca tgctgagact tggggcaggc acacatactg ctgcatgaat tcggagtacg   12840

tgatgtgtgc ccgacgctcc actgtgcaac gctcctcctc cattactgct gctgcaattc   12900

ccagcccatt ctgtttcctg cagggacaac aggtcaggcc agtcaaagct ttaagttcag   12960

gctccctccc gccacccatt tcctatggct caccatcccc cttcgccaga ggcaggcaga   13020

atcacagtca ccatcatcca tagtacaaag agcaaaagca gaccgcgccg cccagccgcc   13080

gccatgaaca tacgggtgca ggcgcgcagc acgctgggaa ttctgtcgac acttgcgtaa   13140

ccacaagcca ttggcttatc atccggactg ctgtagctac ctgcaagcag tagccagccg   13200

gccttggaag tctcctcctt cttcatacaa gcgcattgta cacacacaca cacacacaca   13260

cacacatcca cacaacatcc ctatttaatc atttcctcaa ggcggaggga aaattgtatc   13320

gagcagccat tgactactgt gcgttaccat ggctacagcg cccctaccgg aagttcccat   13380

gtgctagaaa gttcccaccc aaggcgcaat gaaagtgaca agtggccaaa cagcagctct   13440

ggattttat tgggttaaaa agttgtagcc accttccatc ctggtacaca ggatgggtaa   13500

ggggctgtga gcctagtccg gtggacatcg ggatcgagct gaggacgctg catcagcctt   13560

agggaaggcc aggtcaggag tattcacaga ggtggccgca gcctgcaggg cagtgggagc   13620

tgccttccag ccatttcatg atgtgctgaa gatggccacc atgactgcaa ccctgacacc   13680

acacgaacag gcctttaacc acatgatggc agacagcaca catgctggca cagcggtggc   13740

acctgggaat gggagggctg ttcagctggg gagctctggg ctgcctccct tccaacctcc   13800

ccaactgcac catcgcacct gtcacagacc cagccgcggc tgctcatggg ccgcttgcag   13860

tgactacagt tgacatgcag cgtggtggag gcttggttga ggcagctaac tgccctgctg   13920

gtgctcagct tcaccacctc attggacaca ttccagaggc agaagcgctg cagcagatct   13980

atgtaggatg tgtaccagtg ctcctggtgg ggagggacag gctcaggtta gtgtttccag   14040

gaacactagg gtgtgtgtgt ggggggggg ggggtacag tggaaaagaa acaggctcag   14100

atgagggaac cagcagccag tctgggggaga ccctcccccct gccccgcccc gatccatcac   14160
```

65

```
acacctgggt ctgctcatcg atgtcctttc gcactctctc acccagcacg atgagaacag   14220

acacagccat ctgcacgtcg ccctgctcag cgtagaagcg caacatgtcc cgtaccaata   14280

cactgaagaa gtcaggtggc aggcggctgt catagagtgc atgtgagacg gagaggaggg   14340

aggaggagga gtccaccggg gctaatgaag ctgtgtctgc ctcattgccg ctcacgtgcg   14400

gagagtccgc cttgtcctgc agatgctcgg gcccagatgg tgtgtctaca atttcatggc   14460

gcagtgggaa ggcctcttgt ggcagcacat attcaggctc ctcggctggg aggcaggggt   14520

gaggaaagat ggacctgagt gtaggcttcc cgctccactc agctcctgct cacagagcca   14580

tgttctgggt tggacccctc tatagctgac ccagccccta ctcacagtgc acatgttctg   14640

tgtcaagtgg gtacagctca tcgtcctctc cttctacgtc acccagcagg tagtcagcag   14700

gcacatcact cccttctgtt tcctcattat ctgcccagca gggaagggag gtcattcagg   14760

gagtggcacc ccagcctttt cccatcagca tccctgccta ccctcattgg tgacgagtgt   14820

ggctgaagag tcaagcagcg ccgcatcact ccgtgtgtcc cctttgctgc gatccagccg   14880

tgtctcactg cccagccccg ggcccatatc cttcagattg aagctatagg aagggaaggc   14940

ttgtgggtgg ccaggccccg gcagtccact ctggggaggg ggcaggaagt agccaaggat   15000

ctacctgttc atgagcggta ggccgcagga gctgcctttg ccgacgctgt ggttgagatt   15060

ggcagaggat accaagccag ggctacagta gataatccgt aacatggtcc atgtctgtgc   15120

cacctagggg gagggaccaa ccactcaggg gtctccatct cctgagcaga gtcatatacc   15180

aacccatggc ctgatctact accacccagg aaactggccc ctggaaagag cagtgtccac   15240

tctacagtag cacatctgtg aggtgtctgg tgctaagatt agcaatggaa cagatatggg   15300

gaagagagag agcagaaggc agctccaaca ccaagaacga ggtagaagag cgggctgtag   15360

gcacgggact ccttggcctc caccctgagt ttgggctagt ctggttccca tgatctttac   15420

actgtcaaat actgtgacag aggctgccac tgtgactgag ctcagggggg ttgtgggcat   15480

gatgccccaa ctacagcagc agtctcggac cctgggtctc cagcttcaca tacctgattg   15540

cggccaagct cccgagccac ttttgcattg tgatcacaga gctcggctag tggccggcca   15600

gccagaacat aacgttcagc tgtatccacg aaccagctca tgctgccacc gctagactcc   15660

gtctcaaaga cgctcagtgc actggaagcg aggccagaaa aaggttctgc agggtccaac   15720

ttgcgcttga agaagatggg gtggcgccgg tcaccagcat agggtttacg tccagattca   15780

gcagccacca ggctttcctt aacagcaaag gccaggtccc caaagaggcc atagcagaga   15840

ccctcagggt tggcacgctc aacaggctga ctggcatccc gaaacaggtg ctggcataac   15900

gtgctgtcct tggagccgga gagcaggaaa gaggggtcat gcggatgacg ccaggcaatg   15960

cctgttgtga catcacgatg ctcttcaaac atagcagctg gcacgaaggg gcggcgtaca   16020
```

66

```
tcccatacat agatgttgtg gtccaccatc atggaacacg tggccaggtg gtggcggcat   16080

tcaggccgcc acttgactcg ggccactgaa gcgattgtct gcacacagtg tatctctttg   16140

gcacgatgtg tggtcatgtc ccagaccttt accatcttgt cacgtccacc cgtggccagc   16200

cagcccctgc ggagaatgtc ccactcaagc tgtcagtgac ctctcccatg ggaaaggaga   16260

ccctgccaca cccaaaggca tcgtttagcc tccagcctta aggcccatct tctccccaca   16320

ccacacctgt cctctgggtg ccaatcacag cagaagacag gcccattgtg ggctgtgaac   16380

atcctctcac agcggtcggg tcggcggatg tcccagagct gtacgttgcc attctcaaag   16440

gtggaagcaa aggtaaagta gtcacggata ctgaactgca catcccgcac gctctcagac   16500

tggcctgtgg gcacacagtg gtacagacat aagtgagggg ctggccccca gggcgtgggg   16560

gagcccttgg agtggctcca accatgttat tcaccagcct atagcccagc ccagaaacac   16620

aggactggag aaacaccacc aaaatctgcc taggccagag tatgccaggt agtttcagtg   16680

cagactccca gcctcagttc cctagattct gcacaggcaa agctgcagac ctgcctcaga   16740

aggctatgtg tgagcactgg aggcacagga gctcaccctc gttggctgtc tatatgtccc   16800

aggccaaggc ttgtggaaga cccaacgaat ggaagaggag gctcctcgtg ctctgctcaa   16860

tgccctgcca gctgcctagt tagcctcgaa ggaccgaaga gctcagagac catacccacc   16920

ttaaaagatc gccaaacaag gagcctgcgg gtacagtgtc cagtgctttt ccctgccaca   16980

gacagctcaa ggtctgttgt aacccccaaa ttctccaggc cccaaggctc tgcctgcgca   17040

tcatttggtc tcattaaaga tctgagagag gggaatgagg gctgaatgca gacagcttgg   17100

cctgccatga aaggggcgct gtgtactccc cacctcactc tgagtcccgg tctgagtcct   17160

gaagacacaa ggctccgcca ctcactacgg tgtgtctgaa ggctccgcct tccctattag   17220

ccacacccac cacccagtct caatttccac tggtctgacc tcctgcctcg ccccccccccc   17280

caccccaaca gcagtgttgg gcttatcctt ctcaggggggc caaccgttcc catttccacc   17340

actctttctg ggctgttccg cctctgaggg tccatctgag ggcctctgat ggacacacat   17400

ttcagcctgc ctgtcaggtc taacacctac agagatccag atgtacccccc aacttcctct   17460

gtgggcctca ccagagaagg tgctgacaga gtccttcctt cggaggtcaa agcacttcat   17520

gaagccatcc tgagagccac tgagcagcac atgggcctcg gtgggatgga aacagacttt   17580

gttcactgtg cgcttgtgtt ctgtgaacag ctgatcctgc ttgttgcggg atggccgacc   17640

taggttccat gtaaccacca caccattggt ggctgctgtg gccagcagat tctcatccat   17700

ctggtgccag accacatcag cacagctcag gttgagcgag ggcttcctac ccacacgtaa   17760

attcagcttc tctacgaact gctcctcctc aatggcataa atcttgaaga tgcttcggcc   17820

tgccacgaca acctgggcag cgtcgcggca cacgctgatg gcattagctg gcgcgtccag   17880
```

                                        67

```
atggcagtgc atggttcggc ccgtcagagc actgccgctc agggctgtgc taactcggga  17940

catcttttcc atgtctgcac aggtatcaga ggcaggagga ggtcagtgaa gtgtgctgac  18000

ctggtcagcc atggtgtcct aacaattcag gttcagtctc ctctcgaacc ttcggcctga  18060

gatggtccat cagccagtcg ttatcttagg aaggttcaca ttcattggag ttctccagaa  18120

ggctccagag ctaacctgag gttcgtcagc accatatctg acgtctcccg attctatgtg  18180

gttaattctc catcagtgag atcctggggg cggggagtgg ctcctgactt cacagcccca  18240

tgcctccgag gattgcgggg cttccaagag gaaaaacaac cgagtggaaa gggcagagaa  18300

gctagatggg tatggactga gttcatacac cttggaatgc agaagagagg ctacagcaat  18360

aagtgcaggt gtctaattgc accgtgcagc ggcaactgag ggaccagaac caggtgctag  18420

aaaagcctct gccccgtcgc acgccagtca tttaccgaaa gcacagagtg gtcttcagct  18480

acagtaccgg aagagactca cctactgccc gcccattacc ggaagtagtc ctttcttatt  18540

ctcacttcgc cggaagtgag tgacaactcg tctttcacac cataagagtt ccgtgtctgc  18600

gaacagcgct tctgcagagc aaacggacca gaattgggct ccctggtttt aacttggtta  18660

tatatgtccg ctattgatcg tttggttagg gtttctgttt tgttttgttt tgttttgttt  18720

tgttttatgg ttctggaaat cgagcccgag agacaagttc tctaccaact aggctatgtt  18780

tattttaatc tcaatatctt aataccagta cttaaacctt ccctccggtc ttggagaagt  18840

aattttcttt ttctttttttt ctttctttct tttctttttct tttttctttt tctttttttt  18900

ttctttttttg gtttttttgtt ttttgtttgt ttgttttttgt tgtttttgttt tgtttttttgt  18960

ttttcgagac agggtttctt tgtacatccc tggctgtact ggaattcact ctgtagacca  19020

ggctggcctc gaactcagaa atccgtactg ggattaaagg cgtgcaccac cactgcccgg  19080

cgaggagaaa ttttctaagg agaccttgag ttccaccact ctgccacagt gccacttccc  19140

ctacggaagc actagcagcc tggtgccttt tctactctag gtcattattt cactgtgtat  19200

aggtctcaga gacctgaccc tgacctttct ggctcatctg cagctctacc tttcttggct  19260

cctcctttct cagcttgttt ctatcccagc ttccttccct ctgagtattt tccttcttct  19320

ttcacagttc atcagctgtg ttataaaacc atggcttcta cctatggctt ctacccctgc  19380

agaaggtgtc tcctttacac ataggtaggc acttagctca aatgcctacc atcttaggcc  19440

atccttgcca ccccagtcta atgtctcctt tgtctactgt ggtttgtatg attacacctg  19500

ttttggtact ttcagagcct gacaaccagt catcagtggc ttagtagatt tagtctccat  19560

tgtaagaaat cagtgttggc tgaggagcct caagaccaag ttctcagcat aaagaaagat  19620

tgtttgcccc aaagggacaa aggctaggca ataggggacag acagcagata gagggtcagg  19680

gaggagaggg agtagggata tttgaccaga ggggataaag aactgcctct ggatagaaag  19740
```

68

gagacagaca tggcctatag acaaatgaca ctttatgact ttataagtgt aagagggaga   19800

aacaccatgt tagaatgagg tatttaattt tgattagaca ggttaattaa gtgaacccat   19860

gggggctttt gattgctgga cctcaatact ttgatagctg ggccttcatc gccagcctca   19920

ggagtagaag tggccaaata agggaacata ccatggtggc tagctctagg aatataatct   19980

aaggttttta gcaagacaga aggcattgca gagaaggaca aagctggcta gcagtcctcc   20040

cttcagcact tgccctttct cttgaggtgg acacaatagt cttgatcttc tagcagtctg   20100

tgattattta tttatttaat cttttatttga gtgtgtgcac atgtaaatgc aaaaatgtgc   20160

ctgtgtgagt ttttatttgt accaatgtgt gttcaggttc cctaggcggc cagagaacat   20220

catatcccta aagctagagt tactgggcag cttaaaggca tctgacgtga gtgctgggaa   20280

actgaccсca ggtcttctgt aggaacagta gattctttgc actaccgagc catctatcca   20340

gcatcactgg attattttag aacatcacaa tccctcactt agtgaccccg tgcagaagat   20400

aaggctcatt tctgctcttg gatagcaaga tcctagatcc ttgttttctt tttctttttt   20460

ttttctttt tctttttttt tctttttttct ttttctttt cttttctttt tttttctga   20520

gacagggttt ctctgtgtag ttctggctgt cctggaactc actctgtaga ccaggatgac   20580

cttgaactca gaaatccgcc tgcctctgtc tcccaagtgc tgggattaaa ggtgtgcgcc   20640

accaccgccc ctagatcctt tggaagacct acattttgac atcaaacaca ttgttccctt   20700

ttgtgcattc aggacattag agaacttatg ttcgtcagtt gctaaggttc cagtccagcc   20760

ctgagcctcc aacctctcct gtttatggag tacctctaat gatttaggag gcaggaagac   20820

ccagagacct gggccatcca gggcagagtt atcccttcca acttggtcag gtgaccaaaa   20880

aataataacc cacttgtgtt tttacaactt ttttaatata tatttttata aacaggtcac   20940

gtgataaaat agcacaagaa acacttacca aatataaggt tatatcttcc gcatatacag   21000

gagaatgagg tcgttatgta caataagaaa atgattttag gggttggttg gttttgtttg   21060

tgttcctctc tccccttaaa ttttcctcct acagtcattg gaaatatcac agcttcagtt   21120

gcattaatac tttgggcaaa tggacagctt ccccctccct gttgggggac tgtgggaaga   21180

ggggctgaca aatctggctc ctgaccactt cagcctgggg cctccctggc caacactgca   21240

ggggcaggga actctttggg ctgttgaatc tttacttttc gtcaacagca tctctctctc   21300

tcctctctct ctctctctct ctctctctct ctctctctct ctctctctcc ctctggctct   21360

ctgggaatta gtttactctc ttccaaccag ataggggtgt cccaagatta gttgtgggtg   21420

ctgtgtctcc tggggttcaa ggttgggagg gaaggttccc agtcccttcc tcggtattgt   21480

cacctatatg ccaggatccc ctaggacttc caaatttctt gagacctgac agaacatggt   21540

cacaaaccct ttgggcaggc aaaggttttg gaagctggga gaactaaagg aggactggga   21600

```
aggagccctc cttccctggg gtgtgaaggg ggtgctggcc actgggagat cagaatcgct   21660

ggtcagtact gaccctgact catggccagg ccctctgcag attccacctg aggagaccta   21720

tccagggaga agcacagtgc tgagaggcct gagctcagct ggggtagggg ctcaaatggg   21780

tgggctgatc ccaaggcctg atagccctgt gggtggctcc agattggggt ctgccaggtg   21840

cctattccac ccttgcactt aggtggatgg gcaccggtta ggctaatgtt gattgttcca   21900

aggtctgcgt cctcctaaag ctgtcccaac ctggccaggg cccaaggatc tgtaggccag   21960

ggggaaagaa ggagatacgg gggtaggggc gcaggataca gaggacaaga actgacctat   22020

atggcaacaa ggaagaaggg aagccactag gtcaggtggg catggggagt tgggctcctc   22080

aggggggccat ggctttggca gtggcctgtg gcgcctcata ccctttggag ggctccaggc   22140

gcggtgcagt ggaagatgtg ggaggccctt gggggctctt ccaggcccct aagtcctggg   22200

ggaccaatcc ctccatcccc ttgggccctg cccgggatgg ggggacagcg aggggaggca   22260

gagagcagaa tgaagtgggg aagccagtga gggcatatgc actcaaagca tccccaaaag   22320

gacctagacc cccagaaggc gggcgagtga tggcggcggc tgtggcatgg tggagggcgg   22380

gcgtgcagtg cgggcacgag gggcatgcac aaagtccccg agtgtgcgtg cgtgcgtggg   22440

gctggcccct ctgcctcttg ccccagacag ggaaggagca gttgagcaga ggagcaaggg   22500

aggcacgggg cgggatttcc ctcagcttgc cccgcctcca gcgcagggcg gagaagagat   22560

gctccggcag ggatggaaga agggctactg cccagggcgt ggccggttcc cgcggccggg   22620

tcgggaagcc tgcgctattc aatgacgagg cagcgcggca ggtgctgcga gaagtacttg   22680

aagagctcgg gtgtggctcc agggcagttg gtcagctcca gctcctccag ctcttgcagt   22740

tgcacgaggc cagatagccc ggtggtagtt aggagcgggc agcctgcaag catgaggcac   22800

ggggacagcc gctcagcagt tcagttccgc ctggcgggag gactggactg ggctgggcag   22860

gtgggaaagg ctgaagagga caatgacctg ctaggggtgg attctactgg gcttgctgaa   22920

tgaggcggga cctgtaccgg gcgtggaggg gtgggaatct gacaggatgg gcgggacctc   22980

caccagggaa ggatgtgttc agacttggcc tgatgcaggg tagatctgag cagggcgtca   23040

ccaacagggc ttggattcac cctctaaccc gctacctggg atgcgatttg ccaaaacctg   23100

aaatgaatct atctgcacaa aggacctgat gggggtgggg gtcagagtga gtgggcatcg   23160

ttgcccatga ggcatgcggg ggtgggggggt gtctcacctg ctagagacaa gagtcgcaaa   23220

tttctcatgg ctaagaggtg cttcaacccg aagtcctgca cctgggcggg agtgagagga   23280

ggtcgtgtta acttatttcc acttgcttgg gacctggagc cactagaggg gcaggaggta   23340

ggggttcaaa atggagattt ctcagtccca cccccacccc ccgtagactg gaagcaccaa   23400

gagggccttg gggaatctgg aacccaatga ttctgcttag tttggctgct ggttctagag   23460
```

70

```
gactctgaat gaggcggggt gaagttcaag gataatctgt actggtcctc caccctgtcc   23520

tgtcccgcca ccaaccacgc tctaatttcc agatccctgg gggcatggag gtacctggca   23580

gcaccatcgc aggtagaggc tgcggagaga cgacatggtg gacaaatagc tgagaccagt   23640

gtccgtgatg cgtacacacc tgtggttgca ccagaagggg gactgcccac ctttaaatca   23700

ctttacccca tccccaccta ggtccttccc ccctttttcag aggtgtccct agagcaagtc   23760

tctgccccac acaccaatgc cctattcttc ccctggtccc accccgtggc ctcccctttt   23820

ggaatatggc ttttcctcca gttttgggca ctctattcta tgtggttttc cccatccccg   23880

aagcctgtga accgggcacc acccactcta ggcttccagt tcctggagat cccaccagta   23940

cccgcctgga gtgtagggat gtgtgcacct gtccagcacg agctcctcca gtcggtgcaa   24000

gtcgcaggcc acatactcca gtgccatatc ggtgatccga gggcaccagg aaaggtcaag   24060

gctgcgcaat ttgcgcaggt tctcagctac aagctccact ccatcatcag tgaccttaga   24120

gcagcctgag aggctgagtg aggtgaggtt gggcagactg tgcaccacat tgaccacccc   24180

atggttggtt atctcccagc aggaaagcag gcgcagcgtg tgggtgctgt ggccctggcg   24240

tgctgtgaag taggccaggg ccgtgtcagt cacgtgatag gcctgaaggc ttagctctgc   24300

caggttgggt agaagttgtg agatggcagc gattgcatca tctgccacat tgatgcagtc   24360

actgacgctc agggaggtga tacgggcact gaggctggac cacaggccag cctcggtgaa   24420

atcgttgcag cctgacagct ccagacgcac cactccctgc atctgctcca gcatcaccta   24480

caggcaggta caaatcagac taagtcagct ccctggctca cgcctgcccg aagaagtttc   24540

tgggattctg gaagcggtgg tctctcaggc gtgggcttgt gcaagcatgc atcaggctga   24600

actccaagga atgcacttaa actgcttgtt atactttccc aatattccaa ggggggaggc   24660

agggcacagg ggacacaagc cagggtgtgg ccagtgcagg tgtttagaac agatgatttg   24720

aatattaaaa ttacaatcca ggtccagggg tgtggggagg cagggcaggg tgcttctagg   24780

agactcttag ctacccctcc acacaccagc ccagcagccc attcaccaag gcagcttcac   24840

aagcaagcag acaccacat gtaattgact actctcttgc accaaccttg acatcttaat   24900

cacttttttaa caagcccaca ttttcacttt gcattaggcc ccgaaatacc atccttgctg   24960

cttttttatcc tgtatacacc tgagtatctt ccaggcccct caacctccca tgtccataag   25020

gttgcttctg catgttcccc accccctcc ccagacccag acccatctgg tcatagccaa   25080

ttttgggaca cacctctaga ccagcgtcag tgatggtgga gcgcttgaga ctcatggcct   25140

tgactccctt cttagagaga gaatagttgt cgatgaactc gcagatgtct aagtcggaga   25200

caccaactag gcagaagccc tcaaaacctc tagcagcaaa gccctgcagg ttcacaaact   25260

ccttctcgcc tccaggcagc acgttgtaca gctccttggc atgcagcaca ggcgtgaggc   25320
```

```
ccgcccagaa cttgggctgg tagagcacac gccgccaagc cttgcacacc tgagctagta   25380

tgcacttctc acatgctgag aaataccaga aaagcccatt aaggatcttc tcgtctgtgg   25440

ccagtggtgg ccgctccact gggggtccag acactggacc cgaggctggt ccactggctg   25500

gggggcatgg cccccggcc agagcaaccc gggatagtgg tgtagccagg cttggaggtg   25560

ggagggtcgg tgggggtggt ggctggcaag gacggttctt ggcagcaggt gtgcccttgg   25620

tgatgctggc tgcacctagg ccgttgggct ggccaggcag cttcaccaga ccgtttcgag   25680

gcaagcatgg aggcttgggg tcaccatcga taccgggggct cgacattttt cttgcacgct   25740

ctgtggatga gacaggagga acaaataagt gtgccaggct ctggaggaac tggagagcta   25800

acctttctc atcctcgagg ctcctcaccc aaagcatttc caggaccact taacagaaca    25860

gatcctagta tgcgcctttc tctgtctgca caagaggcaa gctggccaca ttcacagcta   25920

gctcaagtgt ccacatcctc tctctgtcct attcactggc tcagcctgcc ttgctcaaag   25980

gtgctttctg atgctctgat tcccatcgag gcttggggag atggaatcct tttttttttt   26040

tttttttttt cgtttttcaa gaagacaggg tttctctgtg tagccctggt tgtcctgtag   26100

accaggctgg ccttgaactc agaaatctgt ctgcctctgc ttcccgagtg ctgggattaa   26160

aggcatgcgc caccatgccc ggtggtagat ggaatcctta actgttcctt tcacctccaa   26220

tcctctgatg cccttcctg gagtttatct tgaggactag gatcggagaa cccagagaag    26280

aaaggtgctt gaatagccct gagtcacaga ccagctctcc aaaatctgtg cacctagtca   26340

gtggtggcac atgccagtgc ttgggaggca gaggcaggca gatctctgag ttcaagagca   26400

gccttgtaca gagagagttc caggacagcc aggactacac agaggagccc aatctcaaaa   26460

gagaaaaaaa atctgtatcc ctataaaagg ctctgaggct cagtcgcatc ccctgtatag   26520

ctgactgtgc taccagcaca gtgatacttg gacacatgct ttttcatatt taaaaaattt   26580

aaaattttct gtgtacgggt gttctcttgc atgtgtacca catacatatc tgtggaattc   26640

agaagaggac atagatctcc tggagctgga gttacaggcc tctaggtcct ctgtaaaagc   26700

agcaggcacc caagcactga ggcacttctc cagcctgctg gacacatgct ctaatccccc   26760

acccgagtca acgcaaacag tgctgtgctg tctctgtaca accaaggaag tagaggccaa   26820

gaccaagaga actatctgac taagctcatg gccgatagtt aagtggtaga gctaaaattc   26880

ggatcagaac tggaggagga aggtcatatg gaaggccagc atgaccaaca tgtctgaatc   26940

actccctccc tgttcccctc agactcccag gcctgtgact gggttgcaga ggacactgtg   27000

agtggacagc acggagggaa agagcagtga gactctcata atgccatgat ctctgtctgc   27060

tctgcattcc agacatcaca cccaccacca cccagcatcc agaaggatgt gattttttcca  27120

tggatttcac ctgcagatga aatgctcaat gcaacactta aaacattcaa aggcaatact   27180
```

72

```
ctggactcca gatggcccct tctataaatg agtgctcttg gctagcctct ggtgagtcca    27240

tttcctcaac atgaaatctg ggaggtgaca ccacaaactc cttcctgtgg gagggatttg    27300

cagtctatgg tagtgtaaat cctgagtccc agagctgacc caagggcaag gcagggtttc    27360

cctttctgtg tgtataggtg atcccaggat gagcaggatg cacacatgaa ttgtcagttt    27420

ttgccctaca atctccaggg agtcaacaaa agcctctacc aaaccttcag tcataccttc    27480

atacctggct ttggtggctc ataaccttgc cagggagcct ggagcaatac caagggctac    27540

atgaagaagc ttcccaagtg gtgggcactg aagtacacag accctcagca gagaaagact    27600

catgttaccc tttaggttga dacccctttc tggagctaca gggccaagga aattggctac    27660

tctttcttat actctggctt gacacaagag aagtagacaa gaactacctt ctgcctttat    27720

tttgcaattc tgtctctatg gttccctata ttgcacatgc aatctgtgct agtttcccac    27780

ctctgggaag ggaaagtccc gggttaagct ctccccagaa catggtcctc atgatgaacc    27840

tagttctgtg tctaaggatg gccattggca gtccaccta aggccctcag tactgtctcc       27900

tgattactcc ctttctcaat atgactctta agacctgggg agtaggccag ctattccta       27960

aaggtttcag accttggctg ctctttgtgc tcaagaacag gcaatagtag ctacaaaaaa    28020

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa agccagtata gcaagtagca atgctaacac    28080

ttgccaaatg cacactggtt caatcacata agaacccgag gctctgaaga gttatagttt    28140

gcatgctacg gtaagggaag tctaaagtgt ggttacccca attccctgga gcaatatagg    28200

gaagctggcc tgttgcggga cagcagcatc catgatggga aaggatgaga ttcttcctta    28260

tgtgctccaa ttcttatcca gtaccacccg cattcctggg aactgtgctg agggtggaga    28320

gacagcccca gcctgagtac cagccatata ctaagtcttg gtctacctct agactgccca    28380

tggaacctca gccactaggc gcagaatcag cgctctcgga accccgcaag cagcaccacg    28440

gacagctccc gcgctcaatg tggtcagccg ggctttctgc accagagaca gtggagcgga    28500

ggaggaatga tcggtggctc cggcaaaccc agctccacag gaagaggcag agggagaagc    28560

cgcccgcttg gctaggatct ggaggaggct gcggaggcgg tgccagggga gaaataatcc    28620

aagagctgtg ccgatttagg ctgaaaatgc tcccagccgc gggtgtgcac gtggagcgcc    28680

tgagttgctg ggggttgggg gggagatgca cgtggatgcc ccgggcatct gcagggctcc    28740

tttctgcccc tctctgcagc tctggagatt cacgtttcca gaagcacacg atcgacccccc   28800

ttcccacctt ccagacttgg gtcacatagg tgaatccagg ccagattcag tgtctaattc    28860

ccccatccag gcccctgtcc ttggtggaag cctcatctcc acaaagtcaa cctcgctttg    28920

tgattagaga ggttctcact tattctttct ctttgacatt aaccaagcct gagaacccca    28980

caggcgtcca aaaatggaaa cagttacacc tgaactcaag ggctgaggct ttcagagaaa    29040
```

73

```
gctaagcaca cactggaaga ctggaacttt cttgccaggt gagcaagagg ttattgaggg  29100

gttctaagca aggatcactc cacctggccc tgccatctca ccagatgatg acataaggtc  29160

tgctcccagg tatcgttcta tagcccaggc tggccttgta tttacagcca tcctacctct  29220

ctgggattac aaatatgtac tgccatgcct agctagagct ccttaatgat gctgcaactc  29280

ctgcttccta ggagggggt ggggggatgg atcagtggcc aagtgaccca ggttgtctct  29340

cccttcacgg acaagcacct caggagctaa tactgaggag aggccttctg gaatgatgtc  29400

cgaactggca cagcactcag ccccaggccc agtatgtcct gtgcttggtt tggagtagaa  29460

ttctgggtca tctctagact gcaggggagc ttgagatggg actccctagg tccatcctgg  29520

gctaggaagg gacagtgaca acttcctcag ctgctgtgtc ttcatcagcc cgtctcatca  29580

aggtgctctc cgggagcact gactgtagtc tttgccctct ctagaggact cattgtggac  29640

ataaggaaca tggagggtac aatgggtctc ctgggaaagt tccagtactt ggtgatttcc  29700

agattggcat tccacagagt ctgatgctga accacgagga cttgcctctt ttgtatatcc  29760

agtcatcaaa agagtctcct ggcctcatct ccatccattc ctccttccct agccagaccc  29820

ctgctcttca ttctcccacc ccccctcacc tggacaagtg gcggtgccct gcttgtctca  29880

taggctttct cagtgccgag cttgctgcct gcctccatag ataggatccc ctagtactct  29940

gaaaaagcac ctccaggtgt cacctggaga tgtcatcact gtttctaccc gtggcctggt  30000

aactccttga agacaaatag caccatgagt gtgtggcttc tgggtactgc tgaaggaatg  30060

gacaagaaat ggtcacctga accattcagg ttcaggggtg cagctgcccg aggctgagag  30120

ccttggctgg ttggcggggc aggagaagca caagaattta actccgggta tgcttgtggt  30180

cagcatttgt tcaaaagaaa tagattggtg ccagcaccta gctccccaga cccgctgatc  30240

ttgaggtgca ggtactcagg gagaggagat cttttggcta agtccctctt ctggggtttc  30300

atggctaagg aatctgactg gctgaccccc aaaatctgac ctgatcctag gaaaaagaga  30360

caggagtcac caaggagatg ccttcctcca acatccctgc ctcatgactt cctagagcac  30420

caccgttcaa aacaagcctc agaaagggat ggatctttct gatgctgcag agaactgcca  30480

cttgtcccaa gaccccaaag cactcagact aggaacacat ctgttgtacc ttgtgtcttt  30540

ttcttctgcc tcactgggtc ctgtgttttc ctatgatttg actagtggcc tctgcacaca  30600

gctatcttcc agcggggtca ctgctgttac cctgaaatct taaagaccgc caactacaca  30660

ctcaccaatg tctcacccgt catctctctc agtctaggga ctctttccct gagtctgtgc  30720

acaggtcagg atgaaggcca ccttggaggc tggcttctgg cctgcatagc acctgccctg  30780

atcagctcct caggtttgaa aaggacaacc aaaacaccca tctccagttc agtctgtagt  30840

ggcagccccg tggcctgggg tgacgaccag ggtatctgag attgggggca cagagcaggc  30900
```

74

```
aacatgttct gtgacagtgc tgtgactcag gcacattcac atccctgcca ctttgcacgt   30960

gtcaaccaca gcacttttct tccacccccт gctggagctg cctctggggt gtgtgtgtgc   31020

cagggacttg gggatagggа gacccttggc ccaggaggta aacagaccca ggttgaacct   31080

aggtcaatgc tccagaagat caaactaatg ctctgtgacc aaactaaggg caaatagtca   31140

cgacaggctt catcttatca cgcccccccт cccgctgttc tgtagaatat gagaccttca   31200

aaacccaagg ctcctttaga cacctgccaa gtgaggccca cctgcactct ctgcagccag   31260

tttggccatc tgaccactaa tggggacagt tcagggctgc tcactgacct cctgaccttg   31320

ctttgctgcc ttcccccccc cccccagtg ccagtcccgg aaggttcatc agtgactctc   31380

agaggtgaag gcttatggtc aggaaggttg agtgatgccc accagaagcc ccttctagtg   31440

cctctgaaat agagagccca atctctcagg gctacctaga tgcacttgtt ttaacttcat   31500

cagtgcgaga cgggaccтgt tcaggtgtat ttgggagaac aggtgagtgc atacgtgtcc   31560

tggacttaca gtatgggtga ctggggaggc aaatagtgag ggtggcagga aatgaggctg   31620

ctgaatcttg ggctatctca acatcaccca cgttttcaac cttgactttc cctctcttct   31680

cccaatggct cccagcccgg ccctcaggcc acgtctaggc aggaaatcga ggctccacca   31740

tgcctctgac ctgatctagg cagcccctaa tcatccaggc tcctgtccca acaccggccc   31800

cccccccccg cccctcccc ctgctgcgat gaggctgctg tccagcactt catcaaagtg   31860

ttgggctgga acagagggtc ccgtgtcccg aaggggcagc agggaggcct gagaagaccc   31920

tggtcagtgg gccctgggac aagatggcag tggaagtcgg gcctccagga gggagcaggt   31980

gaagtggcat ctgccatctg ctcctgggag gtcagctgct ggctcgggct gggaggggct   32040

agcaggaggg ctgctaggcc agatccaggc ctcaacagtg cagccttctc tcctgggatg   32100

acgggtgggg aggatccccc actgactaac atcctagaag gacatggaag ccctgctcag   32160

caatctcagg atgcccattc cccattttga agccttccaa tcttggccaa ggcccatttt   32220

cttttctag aaccaaagac atgccagccc actccctggt atagagctgt gggcggcagc   32280

tctataccag ctcatgtcca gaaggcgtcc aggacccста ggccacatat ccacatggct   32340

cctccttttt aattgtatca acacctgtca cgtcctccag tccgccggag gctcataggg   32400

ctggagtggg gagagcccag agccaagcca gaaggagtgg gggctgacgg cctgcagttc   32460

cgcgtcctcg cagccaccgt gcttcctctc agcaattccg tggtacggag atacagcctg   32520

accgcgctgc cggtagcgag cggggcgcaa tcaggccggc cactagggtt aggggctgcg   32580

gaaccgctgg ctcagcggcc gaagggtgag gtggggaagg ggctgcgatt acgtaagccc   32640

tgctcaggcg gcggcggccg cggcgcagtc ccaagtcccc accccgcccc ctcctcattg   32700

tagccttcgc ggaggcagtg gctcgacctg gaccctcccc tccccgcagc ctggccgcgc   32760
```

75

```
gcgcacacac accttacctc ggccagggct tctggctcat gccacgctgt ccgccgggcg     32820

acctcgcgtc atggggagct cagcacccga gccgcctagc gcggaggcgg ccgaggtgct     32880

ggactggcta gcagggcccg actgcgcccc gcgcccggcg ccccgcgcca gggccctctc     32940

cgcagagttc cgagcggagc tccccggtgg gagaggatct gccgccctga gcctcccacc     33000

ggggaggctg aggctgtgcc cagataaata aggccgcttt gcagggaacc cggcgggcgc     33060

ctcctccctc cccttcccct cctccctgcg ctcctcctcc gcgctcctcc ccgcgctcct     33120

ccgcggccgc cgccggcgcc gcgcgtgccc cccgcccggc cctaagagcc gcgccctcgc     33180

gcgctcccgg cgtccggacc aatccgggtt ctagagccag gctgaggtgt cgagaatccc     33240

caaccccgcc ccgctcctct ccgcgcgctc ccggtgcggt accgacgggc acgcgccagg     33300

cggcagcacc cgctcagcgg ccggtgttcg cacaataaca aaccccgtag gccgcgcggc     33360

cgaagagcag cttcgagtgt cccatctgca caaagtggtg gcgacccaga atcttgccag     33420

ctgaggagaa ttgcggtcct tggcgcgagc ggttggcagt ggggttccgg atcctgcgca     33480

gatgcaactg gccctatgtg cccttgcttg gattgccacc tgtgagggcg aggaaagaga     33540

agccgagtgg agacatcgac atcgtcgcac ctccacacac acacacacac accctcctta     33600

tggcccctag gattcatatc ccttccgatc ctgttgcgag tcaggcagaa gaatccctgt     33660

tctgggtgac atagtcacac acttgcaagc aagtgactaa agcaagcgtc aatgaaaaca     33720

caagacagtg gagtgatagt ggggaggggt tacttctgac aatgtcacaa atgatctgag     33780

acgccctctc tgtccccata gcacttggta ggaatggttc gagaatcgac ctgccttatg     33840

gagttccaca ggttcgctct gagatgaagt agaggtcagg tggtggcagc ctgtatggac     33900

cagggagggt ggtttgtgta aaaggaccca ttaagggctg ctggagagcc agacctctca     33960

gcaggctaaa gggaccccaa tctgggtctc tcttgagacc agaaacagca gatgttggga     34020

cagcagagta gcgtcttgga agagaaggag cttcagggtg gaatgtcctc agatgagacc     34080

aactgctgtg gagatgggag gtctgtctcc tgggaatgat gtgaggccta ggggcaataa     34140

gaaagggaag taagctgtat cctagctggt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt     34200

gtgtgtgtgt gtgtgtgtgt ggtgtatgat catatatgtg aagggcgaat tcgtttaaac     34260

ctgcaggact agtnnccctt agtgnagggg ttaattnntg nagcctgggc gtataccta      34319
```

<210> 48
<211> 21
<212> DNA
<213> Artificial

<220>
<223> primer, human and rat origin

<400> 48

cacccaccag ccgcactggt c   21

<210> 49
<211> 19
<212> DNA
<213> Artificial

<220>
<223> primer, human and rat origin

<400> 49
gagcagctag ggttcaatg   19

<210> 50
<211> 32
<212> DNA
<213> Artificial

<220>
<223> primer, mouse origin

<400> 50
ggttaattaa ccaccctcca tccccagctt tc   32

<210> 51
<211> 37
<212> DNA
<213> Artificial

<220>
<223> primer, mouse origin

<400> 51
agctttggcc ggcctggtcc ttggacctag gtccttc   37

<210> 52
<211> 32
<212> DNA
<213> Artificial

<220>
<223> primer, mouse origin

<400> 52
aaggcgcgcc cctggaagtc cacagtttct tc   32

<210> 53
<211> 35
<212> DNA
<213> Artificial

<220>
<223> primer, mouse origin

<400> 53
agctttgcga tcgctcagtc caagttcagc aatcg   35

```
<210>    54
<211>    38
<212>    DNA
<213>    Artificial

<220>
<223>    primer, mouse origin

<400>    54
atagtttagc ggccgcccgg taggaaggcg gaattgag                           38


<210>    55
<211>    35
<212>    DNA
<213>    Artificial

<220>
<223>    primer, mouse origin

<400>    55
atagtttagc ggccgccagt tcccttagct ctgag                             35


<210>    56
<211>    20
<212>    DNA
<213>    Artificial

<220>
<223>    primer, mouse origin

<400>    56
tatcacaggt aggtgaccac                                               20


<210>    57
<211>    21
<212>    DNA
<213>    Artificial

<220>
<223>    primer, mouse origin

<400>    57
acagccctgg gggggcaaca c                                             21


<210>    58
<211>    20
<212>    DNA
<213>    Artificial

<220>
<223>    primer, mouse origin

<400>    58
gagccattgc taacggacag                                               20


<210>    59
```

78

```
<211>  20
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  59
cctcatagct ccgaaggatg                                             20


<210>  60
<211>  20
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  60
cagacgttcc caggctgtcg                                             20


<210>  61
<211>  20
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  61
ctacagtaac cgagacctac                                             20


<210>  62
<211>  34
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  62
agctttgttt aaactcaata cagtgcgggt caac                            34


<210>  63
<211>  20
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  63
cgacagcctg ggaacgtctg                                             20


<210>  64
<211>  22
<212>  DNA
```

```
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  64
cctcgaatcg tggatccact ag                                              22


<210>  65
<211>  19
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  65
tccctgcacg accttggtc                                                  19


<210>  66
<211>  20
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  66
gttctaccac ctcatattcg                                                 20


<210>  67
<211>  20
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  67
tttggaaggg aggctgaaag                                                 20


<210>  68
<211>  20
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  68
ggaagaaact gtggacttcc                                                 20


<210>  69
<211>  20
<212>  DNA
<213>  Artificial
```

```
<220>
<223>  primer, mouse origin

<400>  69
gtgaacctag ctctgttgtg                                    20


<210>  70
<211>  20
<212>  DNA
<213>  Artificial

<220>
<223>  primer, mouse origin

<400>  70
accttggtct ctgagactcc                                    20
```

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/01858

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C07K 16/00; A61K 39/395
US CL : 530/387.1, 387.3, 388.8, 388.85, 350, 391.7; 424/130/1, 133.1, 156.1, 181.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 530/387.1, 387.3, 388.8, 388.85, 350, 391.7; 424/130/1, 133.1, 156.1, 181.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 02/05843 (EXELIXIS, INC.) 24 January 2002 (24.01.02), see entire documant, especially abstract.	1-7

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"B"	earlier application or patent published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 May 2003 (13.05.2003)	23 JUN 2003

Name and mailing address of the ISA/US	Authorized officer
Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450	Larry R. Helms
Facsimile No. (703)305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/01858

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
   because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
   because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim Nos.:
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-7

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet(1)) (July 1998)

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claim(s) 1-7, drawn to a method of inhibiting growth of tumor cells by administering an antibody.

Group II, claim(s) 8-12, drawn to a method of screening for agents.

Group III, claim(s) 13, drawn to a method of diagnosing a tumor cell by measuring the RRP expression.

Group IV, claim(s) 14-25, 29-32, drawn to a method of identifying a p53 or p21 modulating agent.

Group V, claim(s) 26-28, drawn to a method of modulating a p53 or p21 cell wherein the modulator binds to a RRP polypeptide and function is restored.

Group VI, claim(s) 33-35, drawn to a method of modulating a p53 or p21 cell wherein the modulator binds to a RRP polypeptide or nucleic acid.

Group VII, claim(s) 36-38, drawn to a method of diagnosing a disease.

Group VIII, claim(s) 39-45, drawn to a nucleic acid, host cell and expression system.

Group IX, claim(s) 46-49, drawn to a transgenic mouse.

Group X, claim(s) 50, drawn to a method of selecting an agent.

Group XI, claim(s) 51, drawn to a method of making an antibody.

The inventions listed as Groups I-XI do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Pursuant to 37 C.F.R. 1.475 (d), the ISA/US considers that where multiple products and processes are claimed, the main invention shall consist of the first invention of the category first mentioned in the claims and the first recited invention of each of the other categories related thereto. Accordingly, the main invention of Group I, comprises the first method of inhibiting the growth of tumor cells by contacting the cells with an antibody. Further pursuant to 37 C.F.R. 1.475 (d), the ISA/US considers that any features which the subsequently recited products and methods share with the main invention does not constitute a special technical feature within the meaning of PCT Rule 13.2 and that each of such products and methods accordingly defines a separate invention.

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

Continuation of B. FIELDS SEARCHED Item 3:
CAPLUS, MEDLINE WEST, USPATFULL, BIOSIS, PIR, SWISSPROT, GENESEQ, GENEMBLE
Search terms: SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, RRP, antibody, rhomboid related protein, humanized, cytotoxic label.